Naval Submarine Medical Research Laboratory

TECHNICAL REPORT 1226 10 March 2003

EFFECTS OF THREE WATCHSTANDING SCHEDULES ON SUBMARINER PHYSIOLOGY, PERFORMANCE AND MOOD

by

James C. MILLER, Ph.D., CPE[1],
LT Jeff DYCHE, MSC, USNR[2],
Rebecca CARDENAS, MS[3], and
LT Walter CARR, MSC, USNR[4]

Released by:
CAPT G. A. Higgins, MSC, USN
Commanding Officer
NavSubMedRschLab

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
10 MAR 2003	**Final**	**01 Oct 2000 - 30 Sep 2001**

4. TITLE AND SUBTITLE
Effects Of Three Watchstanding Schedules On Submariner Physiology, Performance and Mood

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFRL/HEPM, Brooks City-Base, TX 78235; NSMRL, Box 900, Groton, CT 06349-5900; NTI, Inc.,Brooks City-Base, TX 78235

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
The general rationale for this study was based upon issues of manpower, especially quality of life and personnel retention, and submarine system performance. Nine male submariners participated as research subjects, two from fast attack submarines and seven from fleet ballistic missile submarines. All participants were submarine qualified and had watchstanding experience on the 18-hour (12-and-6) schedule. The group of participants experienced three watchstanding schedules: the 18-hour Submarine watch schedule; the traditional, Maritime watch schedule; and the Alternative, compressed watch schedule (6-on, 6-off, 6-on, 12-off, 6-on, 6-off-6-on, 24-off). More was learned from the physiological measures employed in this study, especially polysomnography, than from the performance or subjective measures. After combining the results of this study with information available in the research literature about submarine watchstanding schedules dating as far back as 1949, the following schedules were recommended for further study in sea trials: the A Schedule (compressed-6); a fixed, dogged-6 schedule; and the fixed, 8-hour watch schedule.

15. SUBJECT TERMS
Submarine; watchstanding; schedules; sleep; circadian rhythms; 3-shift; 3-team

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**94**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

EFFECTS OF THREE WATCHSTANDING SCHEDULES ON SUBMARINER PHYSIOLOGY, PERFORMANCE AND MOOD

by

James C. MILLER, Ph.D., CPE[1],
LT Jeff DYCHE, MSC, USNR[2],
Rebecca CARDENAS, MS[3], and
LT Walter CARR, MSC, USNR[4]

Approved and Released by:

CAPT G. A. HIGGINS, MSC, USN
Commanding Officer
NavSubMedRschLab

SUMMARY

The general rationale for the study was based upon issues of manpower, especially quality of life and personnel retention, and submarine system performance. The results may also be generalized to applications other than submarine operations. For example, The Air Force Surgeon General had listed aircrew fatigue as a very high priority issue and, throughout the preceding decade, the National Transportation Safety Board (NTSB) had listed fatigue as one of its top 10 safety issues.

APPROACH

The investigation was carried out in the Chronobiology and Sleep Lab (CASL) of the Warfighter Fatigue Countermeasures Research and Development Program, Air Force Research Laboratory, Brooks AFB TX. The CASL was a temporal isolation facility dedicated to research on fatigue countermeasures that extend and enhance warfighter cognitive performance and physical endurance.

Nine male submariners participated as research subjects, two from fast attack submarines and seven from ballistic missile submarines. All participants were submarine qualified and had watchstanding experience on the 18-hour (12-and-6) schedule. By all measures used, this group appeared to be a collection of normal, enlisted Navy submariners, aged 21 to 40 yr. None appeared to have obvious clinical problems with depression, anxiety, insomnia or excessive daytime sleepiness. None were extreme “owls” or “larks.”

The group of participants experienced three watchstanding schedules. Thus, the group spent three 8-day periods living in the CASL. There was a 2.5-day training period before the first 8-day condition and there were 6-day recovery periods between conditions. During each 8-day session, we collected data for two contiguous 72-hour cycles on each watchstanding schedules. By that time (144 h, 6 days), we expected to be able to quantify the essential characteristics of the participants' circadian rhythms. Before and after the 6-day experimental period, we acquired 24 h of refresher data and 24 h of recovery data, respectively.

The three watchstanding schedules were:

- The 18-h, Submarine (S) watch schedule (21-27 September 2001). This schedule is the one in use in the submarine service. It used three watch sections working four possible 6-hour watch periods on a 6-on, 12-off rotating schedule.
- The traditional, Maritime (M) watch schedule (7-13 September 2001). The participants represented just one of three possible fixed watch sections available in the maritime system, the one that required the most extreme circadian acclimatization: 00:00-04:00 and 12:00-16:00.
- The Alternative (A) watch schedule (24-30 August 2001). This compressed work schedule used three watch sections working four possible rotating 6-hour watch periods on a 6-on, 6-off, 6-on, 12-off, 6-on, 6-off, 6-on, 24-off schedule.

Based upon the research literature, we assumed that:

- The fixed Maritime watchstanders would take five days to stabilize their circadian rhythms since this work-sleep environment supported rhythm stabilization.
- The participants' circadian rhythms would free run in the 18-Hour condition.
- The Alternate schedule was designed such that the participants' circadian rhythms would remain entrained to the boat's clock.

SPECIFIC HYPOTHESES

24-hour work-rest cycles will produce better entrainment of circadian rhythms in physiology and performance to the 24-hour clock than will an 18-hour work-rest cycle. This hypothesis was supported by our Conclusion 6, that the participants adjusted their body clocks quickly to the fixed work-rest schedule of the M Schedule.

Given the same average amount of time in bed and average time spent on watch per 24 hours, both sleep quality and sleep quantity will be worse in an 18-hour work-rest schedule than in 24-hour work-rest cycles. This hypothesis was supported by our Conclusion 4, that more good-quality sleep was acquired during P2 in the A Schedule than in the other two schedules.

Given the same average amount of time in bed and average time spent on watch per 24 hours, performance and mood will be worse in an 18-hour work-rest schedule than in 24-hour work-rest cycles. This hypothesis was not supported.

Given the same average amount of time in bed and average time spent on watch per 24 hours, both sleep quality and sleep quantity will be worse in a standard maritime work-rest schedule than in an alternative 24-hour work-rest cycle. This hypothesis was supported by our Conclusions 4 and 5, that more good-quality sleep was acquired during P2 in the A Schedule than in the other two schedules, and that the need for Recovery sleep was not an issue following the A Schedule but was definitely an issue following the M Schedule.

Given the same average amount of time in bed and average time spent on watch per 24 hours, performance and mood will be worse in a standard maritime work-rest schedule than in an alternative 24-hour work-rest cycle. The mood portion of this hypothesis was supported by our Conclusion 2, that the malaise predicted to occur as a result of circadian rhythm disorder caused by the M Schedule apparently surfaced as a perception detected by the Mood 2-R “Fatigue” scale. The performance portion of the hypothesis was not supported.

RECOMMENDATIONS

After combining the results of this study with information available in the research literature about submarine watchstanding schedules, dating as far back as 1949, the following recommendations were made. If, in fact, the work compression and expansion of time off that may be achieved by lengthening the watch is desirable, then the following 3-team schedules should be considered for sea trials:

- The A Schedule (compressed-6)
- A fixed, dogged-6 schedule

- The fixed, 8-h watch schedule

For 24-h operations in geographically-confined, limited-crew-number situations consider:

- The A Schedule (compressed-6)
- A fixed version of the A Schedule
- A fixed, dogged-6 schedule
- The fixed, 8-h watch schedule
- The fixed, 12-and-12 schedule

If work compression and expansion of time off is to be achieved by shortening the watch, then the following schedule should be considered for sea trials and for 24-h operations in geographically-confined, limited-crew-number situations:

- The fixed, close watch schedule

One or more of the intake tools used in this study may be useful in the selection process for shift and night workers. Morning types may report low satisfaction with night work and may opt out of shift and night work. Conversely, evening types may tend more toward acceptance of night and shift work and thus may be the people who often go on to develop the kinds of health problems generally associated with night and shift work. Thus, the morningness-eveningness questionnaire may provide added value in the selection process by predicting attrition and/or health problems.

PREFACE

This report covers the project period of 1 October 1999 to 30 September 2001. The work was performed under AFRL Job Order Number 71845901 and NSMRL Work Unit 5907. The work was performed by AFRL/HEPM, NSMRL and the following AFRL contractors: NTI, Inc. and Veridian, Inc. The AFRL project manager was Dr. William F. Storm, Chief, Flight Motion Effects Branch (AFRL/HEPM), Biodynamics and Protection Division, Air Force Research Laboratory.

NOTES:
[1]Dr. Miller works at the AFRL/HEPM, Brooks City-Base, TX..
[2]LT Jeff Dyche is currently stationed at NSMRL, Groton, CT.
[3]Ms. Rebecca Cardenas, MS, is a contracted employee for NTI, Inc. at Brooks City-Base, TX.
[4]LT Walter Carr was stationed at NSMRL, Groton, CT. His current duty station is at the Naval Health Research Center, San Diego, CA.

CONTENTS

FIGURES

TABLES

EFFECTS OF THREE WATCHSTANDING SCHEDULES ON SUBMARINER PHYSIOLOGY, PERFORMANCE AND MOOD

INTRODUCTION

OBJECTIVE

The Commander, Submarine Group Two, advised the Chief of Naval Research and Commander, Submarine Force, US Atlantic Fleet, that "aligning the submarine watchstanding cycle with the human body's wake/sleep cycle may lower watchstanding fatigue and enhance individual performance." The investigation reported here assessed this possibility.

Also, in Expeditionary Force deployments, the Air Force faced 2- and 3-shift operations to man ground stations and flight lines. In industry, when 24/7 operations must be covered and 8 or 12 hours per day of work are appropriate, then 4 crews are used. That way, one crew is always taking days off. However, the AF was limited to the use of a 2- or 3-crew solution for 2- and 3-shift ground operations in deployments. We did not have the manpower available to allow one crew to take days off. The traditional maritime watch schedule is a 3-shift, 3-crew solution with 8 h of work (*i.e.,* watch) per 24 hours. We assessed the effectiveness of this shiftwork schedule and two alternatives, one 2-shift and one 3-shift. Thus, we expected that lessons learned from this research effort would be directly applicable to the scheduling of AF personnel in 2- and 3-shift, 3-crew operations in combat.

BACKGROUND

The general rationale for the study was based upon issues of manpower, especially quality of life and personnel retention, and submarine system performance. The results may also be generalized to applications other than submarine operations. For example, the Air Force Surgeon General had listed aircrew fatigue as a very high priority issue and, throughout the preceding decade, the National Transportation Safety Board (NTSB) had listed fatigue as one of its top 10 safety issues.

Given a dwindling number of people available to replace technical personnel on active duty in all service branches, the retention of current personnel had become critical. Quality of life was an important factor associated with retention. Crew fatigue and the equity and predictability of work-rest schedules are very important factors in the quality of life at work. In addition to quality of life issues, impending changes in the Navy submarine operational environment drove the need for this study. Shallow-water operational demands included mine threats, navigation in confined waters, more contacts, and increased amounts of spurious contact data. Lean manning was resulting in more tasks and more responsibilities with less redundancy and supervision per person. Sophisticated sensors and equipment were generating increased information processing demands for the submariner. These work demands required that the submariner perform at his maximum ability throughout each watchstanding period.

To implement a change in the work/rest schedule of submariners, a proposed schedule must benefit, or at least not impair, the operational performance of the submarine and the crew. To collect appropriate data aboard a submarine requires re-coding of ship software or bringing new pieces of hardware and several people aboard. This was viewed as being cumbersome and intrusive, and as an imprudent use of Department of Defense resources. Thus, this initial assessment of a proposed change to the watchstanding schedule was performed in a land-based laboratory. The laboratory approach was recognized as being artificial, but the laboratory setting also afforded a clear test of the concept and provided adequate generalization to submarine operations. The laboratory results were to be assessed later in a sea trial.

SCIENTIFIC RATIONALE

Navy fast-attack submarines had been operating on an 18-hour work-rest cycle since the 1960s. All watchstanders (about 70% of the crew) stood watch for six hours and were then off for 12 hours. Then they repeated the cycle. This practice, coupled with the absence of strong photic time cues (daylight-darkness *Zeitgebers*), caused the circadian rhythms of the watchstanders to dysnchronize from the 24-h daily cycle (Naitoh *et al.*, 1983) and, in some cases, free run with a period of about 24.5 hours instead of entraining to the 24-hour clock (Kelly *et al.*, 1996). Also, some watchstanders developed 18-h cycles in addition to 24-h cycles (Schaefer *et al.*, 1979). In the meantime, the boat's operations and social cycles were maintained on 24-hour Greenwich Mean Time (Zulu) and local time zone clocks, respectively. Thus, compared to the 24-hour cycles existing on the boat, the watchstanders experienced a circadian phase lag that accumulated at about 0.5 hours per day. Like circadian rhythm disorder , this phase lag reportedly induced malaise among watchstanders. The inability of the watchstanders' circadian rhythms to entrain to an 18-hour work-rest cycle was not surprising. Early work on circadian rhythms indicated that the limits of entrainment are about 23 to 27 hours (Wever, 1979; Wever *et al.*, 1983).

On the 18-hour work-rest cycle, watchstanders slept about seven hours per 24 hours (Kelly *et al.*, 1996). Since most people operate best on more than seven hours of sleep per 24 hours (Miller *et al.*, 1999; Williams *et al.*, 1974; also, recommendation by the National Sleep Foundation), many submariners operated with a chronic sleep deficit. Anecdotal reports indicated that the prevalence of both jet-lag-like malaise and sleep deficit were high among submarine watchstanders who worked the 18-hour work-rest cycle. The Naval Submarine Medical Research Laboratory conducted a systematic survey of these issues that will be published in a separate report.

An older set of relevant work-rest investigations by Colquhoun, Blake, Edwards, and Hockey from 1968-69 was reviewed by Hockey and Colquhoun (1972), by Colquhoun (1980) and by Colquhoun *et al.* (1978). They compared rotating and fixed 4-hour watches, 8-hour fixed work periods, and 12-hour fixed work periods across 12 days. The rotating 4-hour system was composed of six 4-hour watches (00-0400, 08-1200, 20-0000, 04-0800, 16-2000, and 1230-1630) across a period of 72 hours (12 male participants). These rotating watches were compared in terms of physiology and performance to fixed watches that occurred at 00-0400 and 1230-1630 each 24 hours (16 male participants).

While the emphasis of these studies was to describe the nature of circadian variations in body temperature and performance, several findings were quite relevant to the proposed investigation. Circadian variations in body temperature and most performance measures correlated positively with each other. The exception was short-term memory. Various tests of short-term memory displayed relatively flat patterns or even an inverted circadian pattern compared to temperature.

Vigilance performance for the rotating 4-hour watch condition correlated positively with body temperature, with a 13% peak-to-trough range in signal detection proportion and an 8% range in response latency, with no significant variability in false alarms. This pattern suggested that perceptual efficiency improved with higher body temperatures, as opposed to just a lowering of the decision criterion. Body temperature and vigilance performance in the fixed watch system phase shifted together across the first five days of the 12-day experiment, indicating an adjustment of the participants' circadian rhythms to the fixed watch schedule. They concluded the 12 days with an approximate range of 16% peak-to-trough range in signal detection proportion and 8% in response latency, with no significant variability in false alarms.

However, this excellent degree of circadian rhythm adjustment had not been noted in a previous study by Kleitman and Jackson (1950), for the same fixed watch periods. The difference in results was attributed by Colquhoun to sleep quality. In the Colquhoun *et al.* studies, the participants had a "single, long sleep between 0430 and 1130," while in the Kleitman study, the participants split their sleep with a primary period between 0400 and 0800 and a secondary period after the afternoon watch. This finding supported the need for a long, protected sleep period for watchstanders. Additionally, work by both Aschoff *et al.* (1971) and Wever (1983, 1989) had shown that non-photic *Zeitgebers* that include a disciplined, regular sleep-wake cycle entrained the body's circadian rhythms to a 24-hour period.

While the linked circadian rhythms of body temperature and vigilance performance measures continued unabated for rotating watchstanders in the Colquhoun *et al.* studies, the fixed 4-hour watch schedule caused an initial flattening and a phase delay (5 h across 5 d) in body temperature and vigilance performance. This finding argued favorably for the usefulness of a rotating watch schedule rather than a fixed watch schedule if one wishes to avoid the painful period of adjustment to a new, fixed schedule as the boat gets underway.

Colquhoun revisited the maritime watchstanding problem a decade later (Colquhoun *et al.*, 1978, 1979). Seeking normative values for their work, they looked back upon their hourly temperature readings from 59 young, healthy Navy personnel who were not standing watches and were sleeping normally at night (Colquhoun *et al.*, 1968; description and figure in Colquhoun *et al.*, 1978). They fitted the group mean data with 24- and 12-hour sine and cosine curves (harmonic analysis), explaining 99% of the variance in the group mean data with the fundamental (24-h period) and first harmonic (12-h period). The resulting, complex curve was composed of a 24-h-period waveform with acrophase at 17:00 and peak-to-peak amplitude 1.06 deg F. The acrophase of the combined curve occurred at 20:00 and the peak-

to-peak amplitude was 1.20 deg F. The minima of the fundamental and combined curves occurred at 05:00 and 04:00, respectively. This curve was taken to represent the normal, underlying pattern of circadian-plus-circasemidian variation in body temperature. In fact, the publication of this information was a benchmark, initiating widespread consideration by circadian rhythm investigators of the 12-h circasemidian rhythm in body temperature.

In their earlier studies, Colquhoun *et al.* had noted that circadian rhythm flattening had not occurred until the last of four contiguous 72-h cycles used in those investigations. Now, they had the opportunity to collect temperature data (at 3-h intervals) from eight submarine sonarmen during a 48-day cruise. The sonarmen worked a "traditional," 1-in-3, 4-h watch system that repeated every 72 hours as follows from 00:00 of day 1: 4 on, 4 off, 4 on, 8 off, 4 on, 12 off, 4 on, 12 off, 4 on, 8 off, 4 on, 4 off. In this system, an individual stands three 4-h watches on day 1, one on day 2 and two on day 3. There were two each 4-, 8- and 12-h time off periods. This schedule allowed two contiguous 8-h sleep periods each 3 days. This capability may have been the main factor in the selection of this watch schedule, as opposed to the traditional maritime watch schedule (1-in-3; fixed 4 on, 8 off) in which one can never get 8 contiguous hours of sleep.

Harmonic analysis (fixed 24- and 12-h-period cosine fits) was attempted for each of 16 contiguous, 72-h cycles (16 cycles x 3 days/cycle = 48 days) for the 8 sonarmen (16 cycles x 8 sonarmen = 128 data samples). Good fits for the 24-h period, fundamental harmonic were achieved in only 68% of these 128 samples. Thus, circadian rhythm disturbances were certainly present. Amplitude declined slightly in the first several cycles, and then more sharply across the 48-day period. Acrophase drifted slightly later in the first several cycles, and then more sharply, also. However, these trends may have been caused by the fact that the cosine wave fit method became less and less effective across the 48 days.

When the circasemidian curve was added to the fundamental, circadian curve, good curve fits were achieved in 88% of the samples. Amplitude and acrophase for the combined curve remained fairly stable for several cycles, and then decreased and increased, respectively, more sharply across the 48-day period. Again, greater inter-subject inconsistencies, as measured by the combined-curve fit, became greater and greater across the 48 days. Probably, circadian rhythm disruptions worsened gradually across the 48 days. Certainly, the sleeping patterns of the sonarmen changed across the 48 days, with more and more sleep periods being taken in the 08:00-00:00 (submarine "day") period and fewer in the 00:00-08:00 (submarine "night") period.

On the basis of these observations and parallel observations of body temperature cycles in four officers standing fixed, 8-h watches, Colquhoun *et al.* (1978) recommended using a 1-in-3, 8-h fixed watch schedule (also, see the review by Colquhoun, 1985). Alternatively, they recommended a modification of the schedule suggested by Nathaniel Kleitman in 1949 in which the 8 hours of watch are completed within a fixed, 12-h period (close watch). They expected that the fixed periods would allow rapid re-alignment circadian rhythms with the new work-rest schedule in the first week of the cruise, and that the 16-h off period would prevent sleep fragmentation and restriction.

A problem faced by watchstanders when there are conflicts between external *Zeitgebers* and internal pacemakers has been referred to as circadian rhythm disorder (Arendt *et al.*, 2000). This is the same problem faced by shiftworkers and transmeridian travelers (jet lag). Apparently, the disorder is generated by varying combinations of sleep disturbance and hormonal phase disturbances. Often, sleep disturbance is caused when sleep initiation was attempted on the circadian upswing of body temperature while circulating melatonin is low or declining, instead of vice versa. This problem occurs for new night workers and for transoceanic flight crews and passengers. Potential countermeasures for circadian rhythm disorder that would be useful to submarine watchstanders include good chronohygiene, such as a 24-hour work-rest schedule; good sleep hygiene, such as a long, regular, protected sleep period; and sleep aids and stimulants (Arendt *et al.*, 2000). The use of bright light therapy in submarines may run counter to a number of operational concerns. This investigation dealt with some countermeasure aspects of chronohygiene and sleep hygiene.

There was some precedent for a comparison of a maritime 18-hour work-rest cycle (6 hours of watch followed by 12 hours off) to the standard maritime 1-in-3 work-rest cycle (4 hours of watch followed by 8 hours off). Such a comparison of self-reported work times was carried out by a Naval Postgraduate School Masters student (Stolgitis, 1969):

> Stolgitis established that the average time spent on daily watch and work duties was 12.33 hours for the 4/8 schedule, and 11.67 hours for the 6/12 schedule; thus, the two schedules produced almost equal work output. On the average, 5.82 hours of sleep were available out of the 8.67 hours for rest-recreation under the 4/8 schedule, while 8.66 hours for sleep were available out of the 9.67 hours for rest-recreation under the 6/12 schedule. By dividing the average potential daily sleep periods in hours by the average daily rest and recreation periods in hours, Stolgitis obtained an index of Sleep Cycle Efficiency (SCE)… For the 4/8 schedule, Stolgitis found an SCE of 0.67. In other words, 67% of the daily rest-recreation period was used for sleeping. A higher SCE of 0.89 was found for the 6/12 schedule... [C]rews in some nuclear submarines preferred the 6/12 schedule to the traditional 4/8 cycle as they found the 6/12 schedule more comfortable. A particularly desirable feature of the 6/12 schedule was that once in every three nights, the crew[member] has a chance to get an uninterrupted stretch of free time of approximately 10 hours 30 minutes; time enough for long uninterrupted sleep if desired. (Johnson and Naitoh, 1974)

Stolgitis' arithmetic comparison of the dogged 4-h watch schedule and the S Schedule used here did not use reported sleep times. If it had, it was likely that sharp increases in reported sleep times during the third-night, nocturnal sleep periods would have been noted, compared to sleep times reported for the other two sleep periods. This kind of pattern was found for US Coast Guard cutter crewmembers working on the traditional maritime 1-in-3 (4/8) work-rest cycle (Miller *et al.*, 1999). Watchstanding ceased temporarily about every 10 days when the cutter tied up at a dock. Recovery sleep peaked during these periods that did not require watchstanding.

It is common for shiftworkers to prefer those work-rest cycles that provide longer stretches of time off than others (compressed work schedules). For example, many nurses prefer 12-hour shift lengths to 8-hours shift lengths. When one examines the arithmetic of shiftwork

scheduling, it was easy to see why (Miller, 1992). The 12-hour solution to 24/7 operations allows longer periods of time off than the 8-hour solution. It was likely that the attractiveness of 12 hours off, compared to eight hours off, between shifts was what Johnson and Naitoh (1974) referred to as "comfort." Certainly, it would appear that the crews could solve their sleep disruption problem by extending the inter-watch period from 8 to 12 hours. Unfortunately when crews seek that kind of comfort, giving themselves more potential sleep time between watches, circadian physiology becomes disrupted.

At least one simple, acceptable and untried alternative to the traditional maritime watch schedule, the 18-hour submarine work-rest cycle, and other schedules tried in previous decades was available. A "simple" schedule uses a familiar, predictable watch length, such as six hours. An "acceptable" schedule allows entrainment to the 24-hour clock and provides for long, uninterrupted, nocturnal sleep periods. Such a schedule was assessed in the present investigation: a compressed-6-h watch schedule.

HYPOTHESES

- 24-hour work-rest cycles will produce better entrainment of circadian rhythms in physiology and performance to the 24-hour clock than will an 18-hour work-rest cycle.
- Given the same average amount of time in bed and average time spent on watch per 24 hours,
    - Both sleep quality and sleep quantity will be worse in an 18-hour work-rest schedule than in 24-hour work-rest cycles.
    - Performance and mood will be worse in an 18-hour work-rest schedule than in 24-hour work-rest cycles.
- Given the same average amount of time in bed and average time spent on watch per 24 hours,
    - Both sleep quality and sleep quantity will be worse in a standard maritime work-rest schedule than in an alternative 24-hour work-rest cycle.
    - Performance and mood will be worse in a standard maritime work-rest schedule than in an alternative 24-hour work-rest cycle.

METHODS

EXPERIMENTAL DESIGN

The experimental design was nested-factorial with repeated measures across three levels of factor A (Schedule) and across two levels of factor B (time), the first and second 72-periods of measurement (Periods 1 and 2). The names of the three Schedule levels (conditions) were Alternative, Maritime and Submarine (18-Hour). Since the participants were submariners, we assumed that a mean deviation outside of approximately 95% confidence limits (2 sdu) would be required before an effect could be judged to be meaningful. The participants were already highly selected and trained to deal with fatigue. Thus, a decrement of less than 2 sdu would not be meaningful--*i.e.,* their performance would still be within acceptable operational limits. Thus, the experiment was designed to be sensitive to a two-standard-deviation effect size for a two-tailed test at a confidence level of (alpha = 0.01; 1 - alpha =) 99% and a power of (1 - beta =) 98% (Cohen, 1988, table 2.3.2, formula 12.2.1) when r = 0.30 for repeated measures within Factor A (*ibid.*, formula 2.3.9). This design required a sample size of 9.

To provide for an acceptable and applicable experimental design, we constrained the investigation as follows:

- Consider only 2- and 3-shift, 3-crew solutions to 24/7 operations.
- Provide identical average daily work demands for all participants.
- Limit photic *Zeitgebers* to two: approximately 100 lux during work and less than 10 lux during rest. Exclude all natural daylight from the facility.
- Use local time (Central time zone) as a non-photic *Zeitgeber.* While there was some variation in their timekeeping practices, most submarines set their clocks to local time in their patrol area. They referred to this clock time so that they could use reduced levels of white lighting in the control room between local sunset and sunrise to facilitate dark adaptation for periscope watchstanding.
- Use a specific schedule for the timing and length of times spent in bed for each work-rest schedule, with equal amounts of time in bed per 24 hours across work-rest schedules.
- Assess the 18-hour submarine crew duty day and the traditional maritime watch schedule as two of the experimental conditions.

The participants were exposed to the three conditions in the order Alternative—Maritime--Submarine. The Submarine (18 h) condition was expected to be the most fatiguing. Placing it last precluded the possibility of cumulative fatigue from that condition confounding measurements in either of the other two conditions.

FACILITIES

The investigation was carried out in the Chronobiology and Sleep Lab (CASL) of the Warfighter Fatigue Countermeasures Research and Development Program, Air Force Research Laboratory (AFRL/HEPM), Building. 1192, Brooks AFB TX. The CASL was developed to provide a temporal isolation facility for conducting research, development, test, and evaluation activities on fatigue countermeasures that extend and enhance warfighter

cognitive performance and physical endurance during sustained aerospace operations. Some of its characteristics and capabilities included:

- 2,200 square feet with four temporally-isolated habitats and another nine bunks available in temporally-isolated crew quarters
- Participant monitoring and recording using closed circuit video system
- Facilities for assessing individual and/or group cognitive performance and vigilance
- Access to biochemistry laboratory with high pressure liquid chromatography, molecular biology, chemistry, histology, and image analysis systems
- EEG, polysomnography, actigraphy, blood pressure, HR, and strength measures

PARTICIPANTS

Nine male submariners in the age range 21 to 40 years were recruited from the Navy population. The submariner population did not include females at this time and was not to include them in the foreseeable future. All participants were submarine qualified and had watchstanding experience on the 18 hour schedule. Participant recruiting included an information and question and answer period of at least 24 hours, and an enrollment period during which additional questions could be asked and answered and informed consent was granted and documented.[1]

DURATION OF THE STUDY

We assumed that:

- The 0000-0400/1200-1600 fixed Maritime watchstanders would take five days to stabilize their circadian rhythms since this work-sleep environment supported rhythm stabilization (Hockey and Colquhoun, 1972).
- The participants' circadian rhythms would free run in the 18 h condition (Kelly *et al.*, 1996).
- The Alternate schedule was designed such that the participants' circadian rhythms would remain entrained to the boat's clock.

Thus, we collected data for two concatenated (4 periods x 18 hours, and 3 periods x 24 hours) 72 h cycles on each schedule (Periods 1 and 2). By that time (144 hours, 6 days), we expected to be able to quantify the essential characteristics of the participants' circadian rhythms. Before and after the 6-day experimental period, we acquired 24 h of refresher data and 24 h of recovery data, respectively. The refresher, first 72 h, second 72 h, and recovery data collection periods were labeled P0, P1, P2, and R, respectively. Period R data collection began two hours after recovery sleep ended. All refresher, watchstanding and recovery “days” started and ended at noon. Thus, the group of participants lived in the CASL from noon on a Thursday through noon on a Friday, eight days later.

Each participant participated in all three conditions. Thus, a participant was required to spend three 8-day periods in the CASL. There was a 2.5-day training period (primarily for

[1] Air Force human research protocol no. FBR-2001-16H, with Naval Submarine Medical Research Laboratory IRB human use review concurrence.

repetitions of performance tasks) before the first 8-day condition and there were 6-day recovery periods between conditions. Thus, an individual's participation at Brooks AFB required a commitment of (3 + 3x8 + 2x6 =) 39 days of elapsed time. The participants wore a wrist actigraph (see methods, below) and recorded oral temperatures and work-wake-sleep patterns for 72 hours before and after the time spent at the CASL and during the 6-day recovery periods.

STABILIZATION DAY

The participants reported to the CASL at noon, 24 hours before starting each watch schedule. They performed all of the tasks associated with watch, drill and training periods (below) once that afternoon/evening (1600-2030). They spent the period, 2200-0600, in bed. The following morning was scheduled as personal time.

WATCH SCHEDULES

See Appendix A for expanded discussions of the schedules and Appendix B for schedule details.

Alternative (A) Watch Schedule (24-30 August 2001)

This schedule used three watch sections working four possible 6-hour watch periods on a 6-on, 6-off, 6-on, 12-off, 6-on, 6-off, 6-on, 24-off schedule. One individual cycled through the four possible 6-hour watch periods each 72 hours, during which he stood (4 periods x 6 hours) 24 hours of watch and was allotted (4 periods x 6 hours) 24 hours for sleeping. The participants slept only during specified periods. Psychometric testing was conducted throughout the watch periods, for up to 50 minutes per hour. Psychometric testing was also conducted during 2- to 3-hour periods emulating a ship's drills and training sessions, for up to 20 minutes per half-hour. The periods between drill and sleep and between sleep and training were available to the participants as non-sleep personal time. Food was available every 6 hours as at sea, during the change of watch.

Traditional Maritime (M) Watch Schedule (7-13 September 2001)

The participants represented just one of three possible watch sections available in the Maritime system, the one that required the most extreme circadian acclimatization: 0000-0400 and 1200-1600. They always had the same 8-hour interval, 1600-0000, available for sleep. In 72 hours, an individual stood (3 periods x 8 hours) 24 hours of watch and was allotted (3 periods x 8 hours) 24 hours for sleeping. Psychometric testing was conducted throughout these watch periods, for up to 50 minutes per hour. Psychometric testing was also conducted during the period 0430-0630, emulating a ship's drill, and 0830-1030, emulating a training session, for up to 30 minutes per half-hour. The periods, 0630-0830 and 1030-1200 were available to the participants as non-sleep personal time. Meals were served at 8-hour intervals, during the change of watch.

18-Hour Submarine (S) Watch Schedule (21-27 September 2001)

This schedule is the one in use in the submarine service. It used three watch sections working four possible 6-hour watch periods on a 6-on, 12-off schedule. One individual cycled through the four possible 6-hour watch periods each 72 hours, during which an individual stood (4 periods x 6 hours) 24 hours of watch and was allotted (4 periods x 6 hours) 24 hours for sleeping. The participants did not sleep *ad lib* as was the practice on submarines at sea. They slept only during the middle six hours of each 12-hour non-watch period. Psychometric testing was conducted throughout the watch periods, for up to 50 minutes per hour. Psychometric testing was also conducted during the 1.5-hour period starting 30 minutes after the end of a watch period, emulating a ship's drill, and during the 1.5-hour period starting 60 minutes after the end of a sleep period, emulating a training session, for up to 20 minutes per half-hour. The periods between drill and sleep and between sleep and training, 0630-0830 and 1030-1200 were available to the participants as non-sleep personal time. Food was available every 6 hours as at sea, during the change of watch.

RECOVERY DAY

The watch schedule ended at noon, 144 h (6 d) after it had begun. The participants remained in the CASL for 24 h of recovery. The afternoon was scheduled as personal time. The participants spent the period, 1800-0600, in bed with the idea that 12 h would allow adequate time for recovery sleep. During the period, 0800-1220, they performed all of the tasks associated with watch, drill and training periods (below).

PREDICTIONS

We used the Department of Defense Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model[2], implemented in a Windows® program, the Fatigue Avoidance Scheduling Tool® (FAST®)[3], to predict relative cognitive effectiveness and acrophase. The model was developed using data from relevant research literature. This investigation served, in part, as one validation effort for the SAFTE model. The acrophase prediction in SAFTE and FAST may be applied to many kinds of, though not all, cognitive functions and also to body temperature.

In the following figures, produced by FAST®, sleep periods are shown as blue bands across the horizontal time line. Work (watch, in this case) periods are shown as red bands.

The vertical axis of the diagram represents composite human performance on a number of relatively simple cognitive tasks, such as mental arithmetic, logical reasoning, etc. These kinds of tasks represent cognitive operations that are required to perform safety-sensitive

---

[2] The SAFTE model had been under development by Dr. Steven Hursh (SAIC and Johns Hopkins University) for more than a decade. In the general architecture of the SAFTE model, a circadian process influences both cognitive effectiveness and sleep regulation. Sleep regulation is dependent upon hours of sleep, hours of wakefulness, current sleep debt, the circadian process and sleep fragmentation (awakenings during a sleep period). Cognitive effectiveness is dependent upon the current balance of the sleep regulation process, the circadian process, and sleep inertia.

[3] Developed by NTI, Inc., Dayton OH, under a Small Business Innovation Research Contract awarded by the Air Force Research Laboratory.

jobs. The axis is scaled from zero to 100%. The expected level of performance effectiveness is based upon the detailed analysis of data from personnel engaged in the performance of these tasks during sleep deprivation studies. The algorithm that creates the prediction has been under development for nearly two decades and represents the most advanced information available at this time.

The oscillating line in the diagram represents expected group average performance on these tasks as determined by time of day, biological rhythms, time spent awake, and amount of sleep. Thus, about half of a group should perform worse than the average predicted by the model. Performance may drop to zero when an individual falls asleep on the job or at the wheel. The likelihood of these happening increases as the predicted average performance gets lower. To place the vertical, performance effectiveness axis in context, we offer the following comments:

- The 90% level is the approximate point at which pilots start admitting that they may not be as competent as they should be for flying duties.
- The 75% level represents effects equivalent to 24 hours of continuous sleep deprivation.
- The 50% level represents effects equivalent to 48 hours of continuous sleep deprivation.

Acrophase is the time each day at which peak body temperature occurs and most kinds of cognitive performance reach their highest levels of the day. The model uses the work-rest cycle to determine the target acrophase for the body clock. The body clock responds slowly, as if with inertia, changing at a rate that is usually one hour or less per day. In the FAST® plots, predicted acrophase is shown as a red line on the graph, referenced to the right-hand, y axis of the plot. The right-hand axis is scaled from midnight at the bottom, through noon to the next midnight at the top of the axis. When the red line drops, phase advance is being predicted. When the red line climbs, phase lag is being predicted.

Our predictions for relative cognitive effectiveness and acrophase are shown for the Alternative (A), Maritime (M) and Submarine (S) watchstanding schedules in Figures 1, 2 and 3, respectively. These predictions assume that the participants slept well and soundly for nearly all of each assigned sleep period. Thus, they represent best-case predictions of performance.

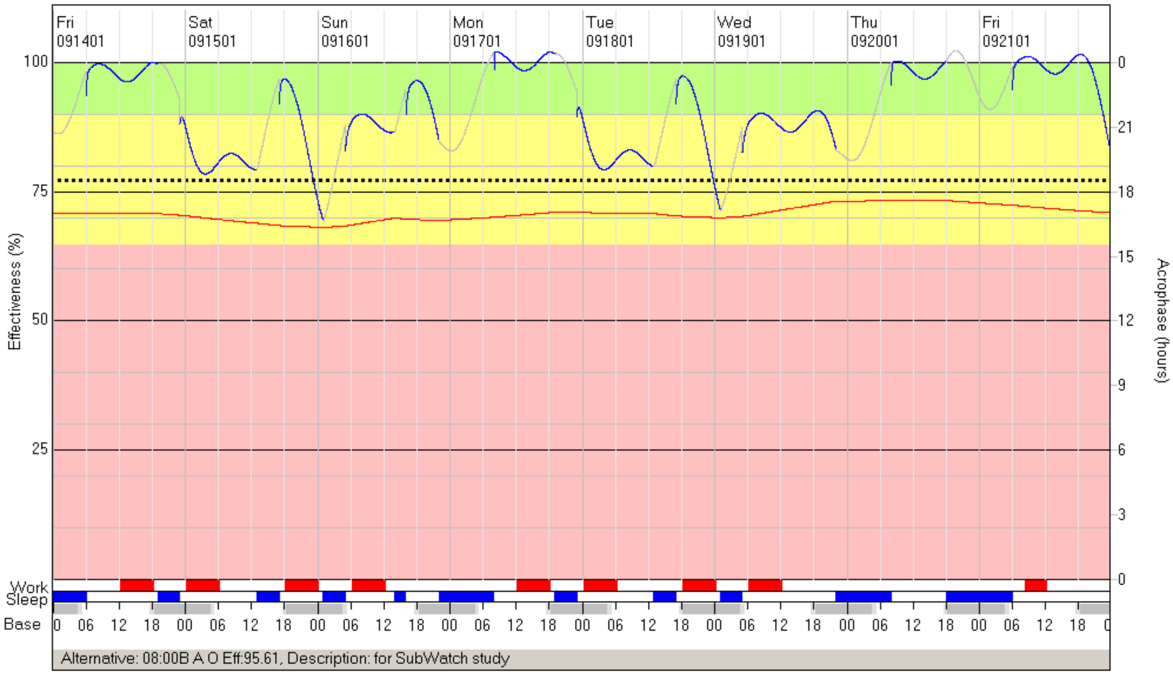

Figure 1. SAFTE/FAST-predictions of relative cognitive effectiveness (blue line) and acrophase (red line) during watchstanding on the Alternative (A) schedule. Near the bottom of the figure, the red bars indicate watch periods and the blue bars indicate sleep periods.

Note that performance during watchstanding on the A schedule was predicted to vary across the six days of watchstanding. Predicted average performance for all watch periods was 89% and ranged from 81-100% for individual watch periods. Predicted average performance for the recovery watch period was 100%. Acrophase was predicted to remain relatively stable.

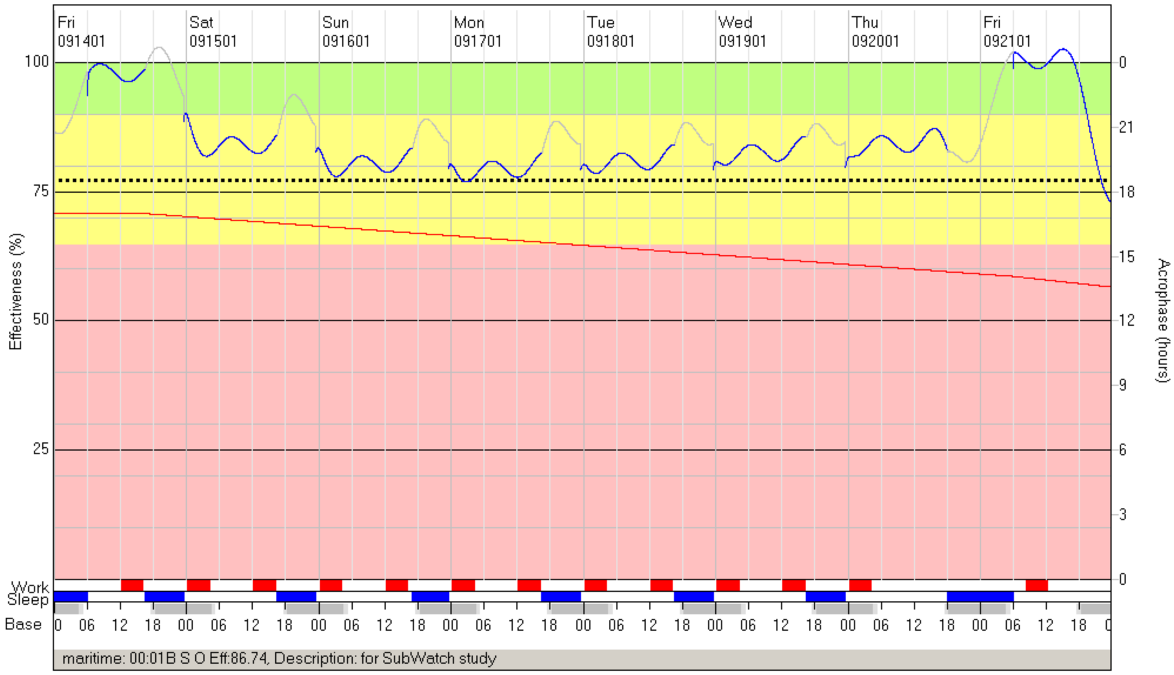

Figure 2. SAFTE/FAST-predictions of relative cognitive effectiveness (blue line) and acrophase (red line) during watchstanding on the Maritime (M) schedule. Near the bottom of the figure, the red bars indicate watch periods (0000-0400 and 1200-1600) and the blue bars indicate sleep periods (1630-2330).

Note that performance during 12-to-4 watchstanding on the M schedule was predicted to drop quickly to the 80% level and increase slowly to the 85% level across the six days of watchstanding. Predicted average performance for all watch periods was 82% and ranged from 78-97% for individual watch periods. Predicted average performance for the recovery watch period was 99%. Acrophase was predicted to phase-advance four hours from 1700 to about 1300, probably producing a mild malaise.

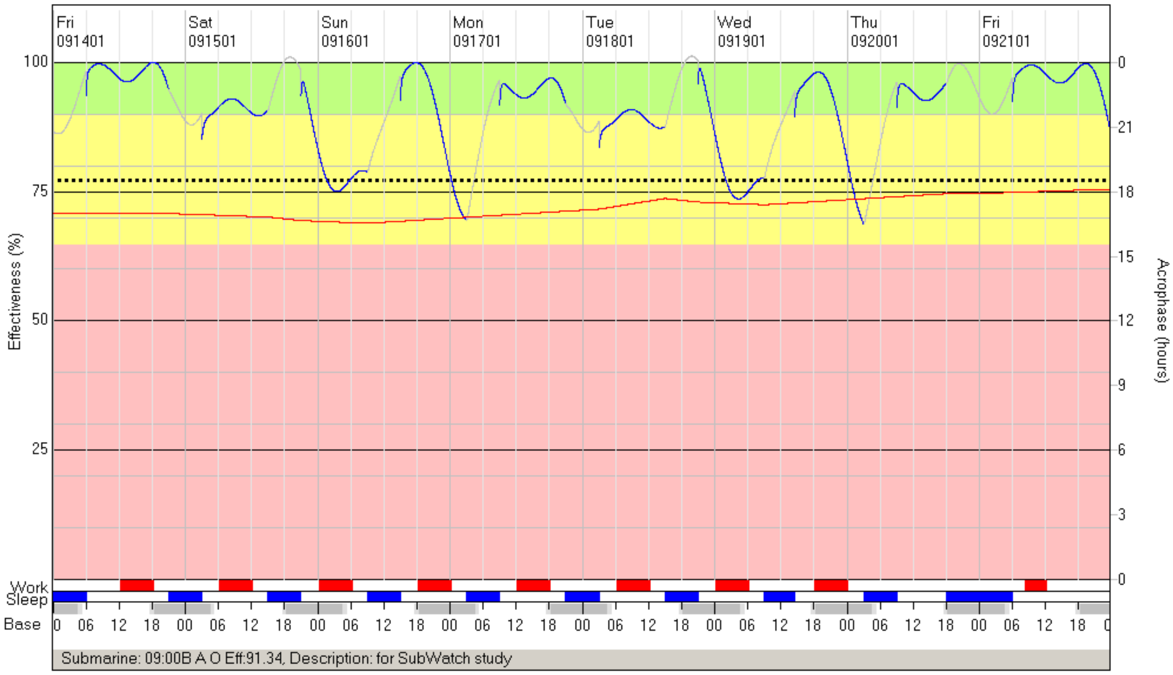

Figure 3. SAFTE/FAST-predictions of relative cognitive effectiveness (blue line) and acrophase (red line) during watchstanding on the Submarine (S) schedule. Near the bottom of the figure, the red bars indicate watch periods and the blue bars indicate sleep periods.

Note that performance during watchstanding on the S schedule was predicted to vary across the six days of watchstanding. Predicted average performance for all watch periods was, as for the A schedule, 89%, and ranged from 77-98% for individual watch periods, much like the M schedule. Predicted average performance for the recovery watch period was 99%. Acrophase was predicted to vary slightly, but then to phase-delay one hour from 1700 to about 1800.

ENVIRONMENT

We limited photic *Zeitgebers* to two levels: approximately 100 lux during work and less than 10 lux during rest, excluding all natural daylight-darkness cues from the facility. We used local time (Central time zone) as a non-photic *Zeitgeber.* Pink noise was broadcast into the CASL at approximately 55 dba. The spectrum and amplitude of the noise were established subjectively by submariner subject matter experts. The temperature and humidity were controlled to be within the range normally encountered on submarines, though the laboratory temperature tended to be somewhat warmer than experienced while submerged. The ambient air in the CASL was not modified beyond normal, household-type air conditioning. Submarine ambient air had specified limits, determined on the basis of preventing human performance impairment. Since normal room air fell within the submarine atmosphere limits, we did not try to simulate a submarine atmosphere. During one sleep period of P1 during the M Schedule, the participants slept experimental testing area due to cooling problems in their normal sleeping quarters.

The participants maintained the general cleanliness of the environment, while a daily cleaning service took care of deep-cleaning tasks. The following recreational equipment was provided or allowed: video movies, computer games, video games, board games, playing cards, dominoes, exercise bicycles, a limited amount of weights, compact disc players, stereo equipment, compact discs. Outside TV and radio broadcasts were not allowed; however, we provided daily print-outs of news summaries from media websites.

POSSIBLE EFFECTS OF A NATIONAL TRAGEDY

The fourth day of watchstanding on the Maritime schedule was 11 September 2001, the morning on which unprecedented terrorist attacks took the lives of thousands in the World Trade Center in New York City, at the Pentagon in Washington DC and in an airliner crash in Pennsylvania. Because of the need to keep our military participants informed of a national crisis that might affect them personally and/or cause the cessation of the investigation and their immediate return to their units, we invoked the following procedures, approximating the methods used to disseminate important information to a submarine crew. We immediately advised the participants orally of the events of the morning as they unfolded. During lunch that day, we provided them with print-outs of relevant news summaries from media websites, and continued that practice throughout the remainder of the investigation. We interrupted testing once that day, briefly, to allow the participants to listen to the initial national radio address given by President Bush.

Subsequently, we found that no direct personal losses had occurred among the participants, that they were not to return immediately to their units and that the investigation had been deemed mission-essential by local USAF commanders, with NSMRL concurrence, and was to continue as planned. The participants were able to view the television presentations of the tragedies during the 6-day break between the Maritime and Submarine watchstanding periods, starting on the afternoon of 14 September. Discussions with the participants brought forth voluntary reports that they were distracted from their cognitive tasks during watchstanding, 1200-1600 on 11 September, and experienced longer-than-usual sleep latencies and difficulty sleeping during the subsequent sleep period, 1630-2330. These problems did not continue after 11 September, according to the participants' anecdotal comments. However, one participant reported continued distraction after learning that his wife's ship, returning from the Persian Gulf at the end of a tour there, had been sent back to the Gulf.[4]

DATA ACQUISITION

Demographic Information

A *Demographics Questionnaire* was used to acquire information about age, height, weight, handedness (preferred), alcohol use, nicotine use, caffeine use, education level, and work

[4] The investigators compliment and thank the participants for their perseverance in the face of this national tragedy. The participants displayed the highest level of professionalism, representing the Navy's submarine service in the best possible manner.

history. In addition, the following instruments were used to acquire measures that could be used as covariates in statistical analyses.

Sleep Behavior Questionnaire. Spielman *et al.* (2000) noted (1) that the use of a sleep history questionnaire should trigger reflection by the participant and, thus, a considered and reliable assessment of his or her sleep behaviors, and (2) that sleep disorders clinics generally use questionnaires of their own design. These questionnaires tend to deal with the adequacy and quality of sleep or to be more comprehensive, including surveys of potential etiological factors. The AFRL Warfighter Fatigue Countermeasures Research and Development Program's (WFC) Sleep Behavior Questionnaire dealt with adequacy and quality. It was designed by JC Miller and PA Hickey (2000). The primary design resources were the questionnaires from The Scripps Research Institute sleep research group, used in Wylie *et al.* (1996), and from Queensland University of Technology (Hubinger, 1998). We approached questions of mild sleep-disturbance etiology through our other questionnaires, below.

Morningness-Eveningness. The Morningness-Eveningness Questionnaire (Horne and Ostberg, 1976) was designed to help reveal tendencies toward morning (lark) or evening (owl) circadian rhythm patterns. These two categories each account for approximately 15% to 20% of the human population and the mid-range category applies to the majority (60% to 70%) of humans. A morning type is defined as one whose circadian rhythms are advanced about two hours (or more) earlier than the norm for the human population as a whole. They awaken naturally between 0400 and 0600 and are ready for sleep by 2000 to 2200. Morning types may be quite sensitive to delays in night sleep and their sleep duration during a morning sleep may be short. Morning types also may report low satisfaction with night work and may opt out of shift and night work (Kundi *et al.*, 1986). Conversely, an evening type is defined as one whose circadian rhythms are delayed about two hours (or more) later than the norm for the population. That is, evening types naturally awaken between 0800 and 1000 and do not feel sleepy until the 0000 to 0200 time frame. Evening types may tend more toward acceptance of night and shift work and thus may be the people who often go on to develop the kinds of health problems generally associated with night and shift work.

Sleep Hygiene and Practices Survey (SHAPS). The SHAPS acquired data concerning participants' knowledge of the effects (1 to 7 scale) of selected daytime behaviors upon sleep and of the presence of caffeine in various OTC medications, food and drink (Lacks and Robert, 1986). Scores on the sleep hygiene knowledge section may range from 13-39; a higher score indicates less sleep hygiene knowledge. Scores on the caffeine knowledge section may range from 0 to 100; a higher score indicates better knowledge of caffeine. Scores on the sleep practices section may range from 0-133; a higher score indicates less healthy sleep hygiene practice.

Beck Depression Inventory. The Beck Depression Inventory is a copyrighted instrument that was introduced in 1961 and revised in 1971. It is a 21-item self-report rating inventory measuring characteristic attitudes and symptoms of depression (Beck *et al.*, 1961). It takes approximately ten minutes to complete and requires a fifth-to-sixth grade reading age. In general terms, a score of 0 to 3 is definitely normal and not depressed; 4 to 7 may be defined as normal or as mildly depressed; 8 to 15 is approximately moderately depressed; 16 to 24 is

definitely clinically depressed; 25 to 30 is severely depressed; 31 and above requires immediate clinical evaluation. We acquired the trait anxiety score as a potential covariate for analyses of sleep quantity and quality data.

Trait Anxiety Inventory. The State-Trait Anxiety Inventory (STAI) is a copyrighted instrument (Spielberger, 1983) designed to differentiate between the temporary condition of "state anxiety" (see below) and the more general and long-standing quality of "trait anxiety" in young adults. The STAI is used widely in sleep disorders centers because it is easily administered and interpreted and it targets a psychopathology that is commonly associated with insomnia (Spielman *et al.*, 2000). The two inventory parts differ in item wording, in response format (intensity vs. frequency), and in the instructions for how to respond. The first self-report inventory consists of 20 items designed to screen young adults for long-standing anxiety problems. Scores may range from 20 to 80. A higher score indicates higher anxiety. The trait anxiety norm for working men, aged 19 to 39 yr, is 35.55 +/- 9.76 (*ibid.*). We acquired the trait anxiety score as a potential covariate for analyses of sleep quantity and quality data.

Cognitive Hardiness Scale. Cognitive hardiness (CH) is a sense of control, commitment to the projects and people in one's life, and a tendency to appraise events as challenges (versus threats). CH appears to moderate the relation between stress and both illness and depression, and has predicted cortisol reactivity. Cognitive hardiness is a construct of stress resiliency, grounded in existential personality theory (Kobasa & Maddi, 1977). Hardiness is composed of three dimensions: commitment, control, and challenge (Kobasa, 1979). Commitment is a tendency to involve oneself in (rather than being alienated from) whatever one does. Committed people have a sense of purpose in their lives, and events and others in their lives are meaningful. Control is a tendency to "feel and act as if one is influential (rather than helpless) in the face of varied contingencies of life" (Kobasa, Maddi, & Kahn, 1982, p. 169). Someone with an internal locus of control is not overwhelmed by life's events, being more likely to transform events into something consistent with one's life plans, and so maintain meaningfulness in one's life. Finally, challenge is the conviction that change, not stability, is the normal life condition. Changes are seen as "interesting incentives to growth rather than threats to security" (Kobasa *et al.*, 1982, p. 170). The Cognitive Hardiness Scale contains 30 belief items, rated from 1 to 5 for agreement. The lowest possible score is 30 and highest possible is 150. A higher score represents greater cognitive hardiness. The population mean is about 106, but the mean in military groups may be 114-118. We acquired the CH score as a potential covariate for analyses of performance data.

Epworth Sleepiness Scale. The Epworth Sleepiness Scale (ESS) was devised at Epworth Hospital in Melbourne Australia (Johns, 1991, 1992). The ESS has correlated well with electroencephalographically- (EEG-) determined sleep latencies measured at night or during the day and is considered to be a validated and reliable self-report measure of sleepiness. The subjects use a number from 0 to 3 corresponding to the likelihood (never, slight, moderate, and high, respectively) that they would fall asleep in eight situations such as sitting and reading, watching TV, as a passenger in a car for an hour, etc. Ratings above 15 out of a possible 24 are cause for concern with respect to acceptable job performance. An ESS rating

was acquired from the participants at intake and then every few days, using a personal log book. We acquired the sleepiness scale rating as a potential covariate for analyses of sleep quantity and quality data.

Periodic Subjective Measures

Stanford Sleepiness Scale. To use the Stanford Sleepiness Scale (SSS; Hoddes *et al.*, 1973), the subject selects one of seven sets of Likert-scale descriptors, ranging from 1, "Feeling active and vital; alert; wide awake," to 7, "Almost in reverie; sleep onset soon; lost struggle to remain awake." The SSS usually correlates with standard measures of performance and usually reflects the effects of sleep loss. However, the extreme values on the scale are used infrequently and the rank-ordered statements overlap several perceptual dimensions including sleepiness-wakefulness, alertness and concentration. Horne (1991) suggested parallelism between the SSS and the alertness-sleepiness descriptors used for the "vigor" factor of the Profile of Mood States (POMS). The POMS vigor scale has also demonstrated sensitivity and reliability with respect to quantifying perceptions of sleepiness. A SSS rating was acquired from the participants every few hours while they were awake, using the personal log book.

Pre-Sleep Arousal Survey. According to Spielman *et al.* (2000), "half-completed thoughts, racing, and the repetition of themes may represent operating characteristics of a mind that is temporarily incapable of thought." "One reason why hyperarousal so effectively forestall sleep is that, once triggered, it takes a long time for baseline conditions… to be re-established." The 12-item, modified Pre-Sleep Arousal Scale (Lacks, 1987; based upon Nicassio *et al.*, 1985) is administered upon awakening from a major sleep period. It assessed, retrospectively, the degree of presence, on a scale of 1 to 5, of a number of perceived cognitive and autonomic-sympathetic symptoms during sleep onset that may have affected sleep latency. Data reduction provided two sums, for cognitive and somatic effects, respectively. We acquired these data to help us determine the etiologies of long sleep latencies.

State Anxiety Inventory. The State Anxiety Inventory is a copyrighted instrument (STAI; Spielberger, 1983) designed to evaluate feelings of apprehension, tension, nervousness, and worry, which increase in response to physical danger and psychological stress. The self-report inventory consists of 20 items. Scores may range from 20 to 80. A higher score indicates higher anxiety. The state anxiety norm for working men, aged 19 to 39 yr, is 36.54 +/- 10.22 (*ibid.*). We acquired the state anxiety score as a potential covariate for analyses of sleep quantity and quality data.

Mental Workload. The 7-point mental workload scale was created by the Crew Performance Branch of the USAF School of Aerospace Medicine in the late 1970s, and then re-examined, linearized, and verified by the Human Factors Branch of the Air Force Flight Test Center (Ames & George, 1993). Individuals choose one of seven sets of statements describing their average mental workload during the preceding work period. The statements range from 1, "Nothing to do; no system demands," to 7. "Overloaded; system unmanageable; essential

tasks undone; unsafe." A mental workload rating was acquired from the participants after watch, drill and training periods using the personal log book.

Physical Workload. The 15-point physical workload scale is one variant of a scale designed to allow estimates of heart rate caused by varying levels of dynamic work (Borg, 1985; Kilbom, 1991). The scale is anchored with a statement at about every other number, and provides a rough estimate (rating x 10) of heart rate for a young, fit male. The statements range from 6, "No exertion at all," to 20, "Maximal exertion." A physical workload rating was acquired from the participants after watch, drill and training periods using the personal log book.

Mood

The Mood Scale 2-R was adapted originally by Thorne (Englund *et al.*, 1987) from the scale of Ryman *et al.* (1974). The scale was shorter than the Profile of Mood States and was selected partly for that reason. It was implemented within the framework of the ANAM library (Reeves *et al.*, 2001). The sub-scales were Activation, Happiness, Depression, Anger, Fatigue, and Fear (anxiety). It consisted of a listing of 36 adjectives. Participants were asked to respond by pressing 1,2,or 3 on the computer keyboard, ("Press 1 for yes, 2 for somewhat, and 3 for no") in response to the question, "How does the word shown below describe how you feel right now."

Symptoms

The WFC Symptoms Checklist was created by JA Gibbons and PA Hickey (2000-2002). The checklist contained 73 items with seven rating levels for each item (none, slight, moderate, or severe). Items were acquired from FDA reports of symptoms associated with sleep aids and alertness aids. Items were also incorporated from the motion sickness symptomatology checklist of Wiker *et al.* (1979).

Physiological Measures

Body Temperature. Oral temperature was taken several times per day while awake. The participants were instructed to refrain from eating or drinking for 15 min prior to scheduled temperature measurements, and proctors monitored this behavior. When participants erred by drinking water in this 15 min period, the temperature measurement was delayed the appropriate number of minutes. We acquired temperature data to estimate the characteristics of the circadian rhythm in body metabolism.

Hormones. Participants aspirated 3-cc saliva samples several times per day while awake. The samples were frozen and then shipped to the analysis site. The samples were analyzed for melatonin and cortisol content by the Endocrine Core Lab, Yerkes Primate Research Center, Emory University (Atlanta, GA).

The Direct Saliva Melatonin RIA kit was used to estimate melatonin by a double-antibody radioimmunoassay based on the Kennaway G280 anti-melatonin antibody (ALPCO

Diagnostics, Windham, NH). Undiluted human saliva samples and reconstituted standards and controls were incubated with the anti-melatonin antibody and ^{125}I melatonin. ^{125}I melatonin competed with melatonin present in samples, standards and controls. After 20 hours of incubation, a solid-phase second antibody was added to the mixture in order to precipitate the antibody bound fraction. After aspiration of the unbound fraction, the antibody bound fraction of ^{125}I melatonin was counted. Results were reported as melatonin (pg/ml). We acquired melatonin data to estimate the characteristics of the circadian rhythm in sleep patterns.

The solid phase enzyme immunoassay for cortisol was a competitive type immunoassay in which horseradish peroxidase-labeled cortisol (HRP-cortisol) competed with cortisol in the sample saliva for a fixed and limited number of antibody sites immobilized on the wells of the microstrips (Diagnostic Systems Laboratories, Webster, TX). Once the competitive immunoreaction has occurred, the wells were rinsed, and the HRP-cortisol fraction bound to the antibody in the solid phase was measured by adding a chromogen/substrate solution that was converted to a blue compound. After 15 minutes of incubation, the enzymatic reaction was stopped with sulfuric acid, which also changed the solution to a yellow color. The absorbance of the solution, measured photometrically at 450nm, was inversely related to the concentration of cortisol present in the sample. Calculation of cortisol content in the sample was made by reference to a calibration curve. Results were reported as Cortisol ug/dl. We acquired cortisol to estimate the characteristics of the circadian rhythm in body metabolism.

Activity. The Actigraph Sleep Watch (Precision Control Design, Inc., Ft Walton Beach FL, available from Ambulatory Monitoring, Inc., Ardsley NY) resembled a wristwatch and was worn in a similar manner. A small accelerometer systematically recorded the individual's movement over time, both while awake and asleep, providing an effective means to identify sleep behavior patterns. The device contained a single-axis piezoelectric accelerometer, a luminance meter, random access memory, an event marker button, selectable filters, and communication connectors. Its battery supported up to 30 days of data collection. The participants were instructed to wear the WAM on the wrist of the non-preferred hand, also removing their watch, if present; and to wear the WAM at all times except while showering or participating in activities where the WAM might impede performance or be subjected to water immersion. The actigraphy data were reduced using the Cole-Kripke sleep scoring algorithm (Cole *et al.*, 1992) to categorize each recorded epoch into sleep and awake periods. We acquired actigraphy data to estimate the sleep patterns of the participants while they lived outside the CASL.

Polysomnography. Sleep quality during the experiment was assessed with ambulatory electrophysiological equipment. The electroecephalogram (EEG) was acquired from the C3-A1 or the C4-A2 scalp leads of the International 10-20 system with an Oxford Medilog ambulatory recorder system and digitized on an Oxford data system (Oxford Instruments Ltd., Abingdon, Oxon, England; seven participants) or an Embla (Flaga HF Medical Devices, Iceland; two participants). The EEG signal was digitized at 128 samples/sec. The EOG and EMG signals were also acquired to support sleep scoring by a registered polysomnographic technologist (LV), using the standardized methods of Rechtschaffen and Kales (1968). The number of electrodes (2 scalp, 1 mastoid, 2 outer canthi, 2 submental) and their methods of attachment (collodion for scalp and mastoid, adhesive rings for outer canthi and submental)

replicated the successful approach we used and reported in Wylie *et al.* (1996). Measures assessed included percents Stage 1, Stage 2, Stage 3, Stage 4, Slow-Wave Sleep (SWS), Stage REM, and Stage Wake (%S1, %S2, %S3, %S4, %SWS, %SREM, %WASO, respectively); sleep latency (to the first three epochs of any sleep stage; SLat); total sleep time (TST); and sleep efficiency (SE; TST / nominal time in bed).

The selection of these measures was related to the structure of the sleep periods embedded in the schedules of the A, M and S watchstanding schedules. Note that 24 h of sleep were scheduled for each 72 h period (an average of 8 h per day) in the A and S schedules. The traditional problem with the M schedule is the inability to acquire 8 h of uninterrupted sleep during the 8 h of between watches. Thus, only 21 h of sleep were scheduled for each 72 h period of the M schedule. These hours were distributed across three 7 h sleep periods per 72 h that occurred each day from 1630 to 2330, between the noon and midnight watches. Subsequently, we were unable to make direct comparisons of absolute numbers of sleep hours across sessions. Thus, to compare sleep qualities across sessions, we relied heavily, though not exclusively, upon percentages for sleep stages and, of course, upon sleep efficiency.

We considered inserting a 3 h nap into the M schedule to bring total time in bed up to 24 h per 72 h Period. The best nap time, biologically, appeared to be 0430 to 0730 during the second 24 h of each 72 h period. With this nap inserted into the simulation shown in Figure 2, the SAFTE model predicted that average performance for all watch periods would increase from 82% to 84% but would still range from 78-97% for individual watch periods. Predicted average performance for the recovery watch period would still be 99%. Instead of phase-advancing monotonically from 1700 to about 1300, acrophase would vary, but still phase-advance two hours from 1700 to about 1500. These mild improvements did not outweigh the advantages of comparing the A and S schedules to an M schedule in which sleep was limited by the length of time off between watches. However, this simulation did show that a structured, planned nap for the 12-to-4 watch would be somewhat beneficial.

Oculometry. The FIT 2500 (PMI, Inc., Rockville MD) pupillography system was used here periodically to help estimate levels of physiological arousal. The system simply required the participant to track visually the apparent motion and flashes of an LED display for about 30 seconds. Minimal training was required and no learning or skill effects were expected. The FIT was designed originally as an industrial fitness-for-duty evaluation system that would detect physiological impairments due to fatigue and many other factors. It was used here simply as an oculomotor tester. We expected to find cumulative fatigue effects expressed as gradual reductions in baseline pupil size and saccade velocity, and increased pupil response latency.

Saccade velocity slowing may (Russo *et al.*, 1999, Rowland *et al.*, 1997, Stampi *et al.*, 1994) or may not (Morris & Miller, 1996) be a useful index of fatigue. Baseline pupil size varies as a function of fatigue and or sleepiness (Pressman *et al.*, 1986; Schmidt *et al.*, 1981; Yoss, 1969; Ranzijn & Lack, 1997). Similarly, increasing pupil response latency may (Russo *et al.*, 1999, Rowland *et al.*, 1997) or may not (Ranzijn & Lack, 1997) be a useful indicator of

fatigue. However, the combining of these three measures has allowed reliable detections of fatigue (personal communications, J. Krichmar and R. Perry, PMI, Inc.).

Grip Strength

Grip strength was measured using the Jamar Hydraulic Hand Dynamometer (Sammon Preston Ltd, Chicago, IL). The participant was seated in a chair without armrests with the shoulder adducted and neutrally rotated, the elbow flexed at 90°, and the forearm and wrist in neutral position (Fess *et al.*, 1984): wrist position between 0° and 30° extension and between 0° and 15° ulnar deviation (Pryce, 1980; Kraft and Detels, 1972). Using the second handle position (Stanley and Tribuzi, 1993; Mathiowetz *et al.*, 1984; MacDermid *et al.* 1994; Fess *et al.*, 1984), the participant squeezed as hard as he could with his dominant hand. As he began to squeeze, the proctor motivated him by saying "harder ... harder ... relax." The participant squeezed three times with a 1-minute rest between trials. We recorded all three scores and assessed the mean and maximum forces generated.

Postural Sway

Postural stability reflected the overall function of visual-vestibular-somatic control systems, integrating somatosensory function, with and without visual function (eyes open and eyes closed), and the static component of vestibular function provided by the otolith organs. Stability was measured using a force platform (model OR6-5-1) and the BEDAS software (both from AMTI, Watertown MA). We used a custom program to batch-process the BEDAS output files. The measure acquired was the 95% confidence ellipse (A95) of the distribution over time of the center of pressure exerted by the participant's feet on the force platform. These measurements had been used previously to analyze postural stability after alcohol ingestion (Kubo *et al.*, 1989), benzodiazepine administration (Patat and Foulhoux, 1985), and prolonged exposure to microgravity in space. The participants stood with heels together, feet open at a 30-deg angle, and hands at their sides, much like a relaxed version of the military position of attention. The total time of a test was two minutes. We announced the elapsed time every 15 or 30 seconds. Two minutes of data were collected for both the eyes open and the eyes closed conditions. The two-minute data collection period was used because, when we looked at cumulative, 30-sec epochs of data, we found that we were unlikely to detect a fatigue-related effect on the A95 measure until all data from a 120-sec test period had been acquired (Eddy *et al.*, 2002).

Performance Measures

Simple Cognitive Performance Battery. A cognitive performance test battery was implemented on desktop personal computers in the Windows® operating system using the Navy's Automated Neuropsychological Assessment Metrics (ANAM) library. It consisted of a library of tests and batteries designed for a broad spectrum of clinical and research applications. This library of computerized tests was constructed to meet the need for measurement of cognitive processing efficiency in a variety of psychological assessment contexts that include neuropsychology, fitness for duty, neurotoxicology, pharmacology, and human factors research (Reeves *et al.*, 2001).

All stimuli were presented on the PC screen, and all performance task responses were made with the PC mouse buttons with the preferred hand. The battery included the following tests.

- Simple response time task: simply required a rapid mouse-button press in response to the display of the asterisk (∗) symbol. There were 20 trials, with an interstimulus interval that varied from 650 to 1100 msec. Timeout (no response) occurred at 1000 msec.
- Mental arithmetic task: required a left or right click corresponding to a < 5 or > 5 solution of an addition-subtraction problem consisting of three single digits. The probe duration was set to 4500ms, with a timeout value of 5000ms. As soon as the subject responded another probe was presented. The task ran for three minutes.
- Delayed matching-to-sample task: required a left or right click corresponding to a left-right choice between two patterns, one of which matched a single pattern presented 5.0 to 5.1 sec previously. The probe duration was set to 3000ms, delay was set to 5000 to 5100ms, and timeout occurred at 3100ms. The task ran for three minutes. The pattern structure was a four-by-four grid, within which eight cells were colored red and eight were colored aqua, in quasi-random patterns.
- Logical reasoning task: required a left or right click corresponding to a true-false choice about a positive or negative statement concerning the order of two symbols. The probe duration was set to 4500ms, with a timeout at 5000ms. As soon as the subject responded another probe was presented.
- Running memory task (one back): required a left or right click corresponding to a true-false choice concerning the identity of a single digit with the preceding digit. Each digit was displayed for 200 msec. The task ran for three minutes with an interstimulus interval that varied from 4.5 to 4.6 sec. Timeout (no response) occurred at 1.5 sec.

Task training on the ANAM was conducted during the 2.5 days immediately preceding the first Schedule. The participants completed the ANAM battery six times during training, twice per day, during those three days. Additionally, they completed a refresher ANAM battery once during the afternoon of the noon-to-noon stabilization period that preceded each of the three 6-day watchstanding periods.

The ANAM test order for watch periods was Simple RT, Mental Arithmetic, Matching to Sample, Logical Reasoning, Running Memory. For drill periods, it was Simple RT (3 times in a row), Running Memory. For training periods, it was Simple RT and then Logical Reasoning. The experimental watch, drill and training periods are explained, below.

Measures acquired from these tasks included percent accuracy, mean response time for correct responses (MNRTC), numbers of omissions, standard deviation of response time for correct responses (SDRTC), and throughput (number correct divided by mean response time for all responses, in units of number correct per minute)

Vigilance Performance. Vigilance performance was assessed using the Psychomotor Vigilance Task (PVT), an extension of the Unprepared Simple Reaction Time Task (Dinges, 1992; Dinges *et al.*, 1997; Vigilance Task Monitor, Model PVT-192, CWE, Inc., Ardmore PA, available from Ambulatory Monitoring, Inc., Ardsley NY). This task is learned quickly (two 1-minute trials) and is sensitive to fatigue due to sleep loss, circadian variation, and shift work. It proved to be sensitive and reliable in field studies of fatigue in commercial truck drivers (Wylie *et al.*, 1996) and US Coast Guard crewmembers (Miller *et al.*, 1999). The 8 x 4.5 x 2.4-inch portable, battery-operated device ran a continuous, simple response time test for ten minutes. The task required sustained attention and discrete motor responses: the participant watched a digital counter (LED) on the device and, when the counter started to run, turned off the counter as quickly as possible with a button press using the preferred hand. The interstimulus interval varied from 2 to 10 sec. A relatively quick response was about 200 msec. Timeout (no response; lapse) occurred at 500 msec.

Task training on the PVT was conducted during the three days immediately preceding the first Schedule. The participants completed the PVT twice during training, once during the afternoon of the first training day and once during the afternoon of the second training day. Additionally, they completed a refresher PVT once during the afternoon of the noon-to-noon stabilization period that preceded each of the three 6-day watchstanding periods.

The variables provided by the PVT-192 included the number of lapses, the mean of the reciprocals of the 10% fastest response times (Mn_FRRT), the mean of the reciprocals of all response times (Mn_RRT), the mean of the reciprocals of the 10% slowest response times (Mn_SRRT), and the standard deviation of all of the response times (SDRT).

Complex Cognitive Performance. Two complex tasks were used, Synthetic Work and SubSkillsNet. SynWin was a Windows® adaptation of the SynWork1 DOS program (Elsmore, 1964) that was created in response to a perceived need for a laboratory performance testing situation intermediate between the tests typical of performance assessment batteries and full-blown simulators or "part" simulators (cf. Alluisi, 1967). In contrast to earlier multi-task performance assessment systems, which typically involved one-of-a-kind hardware devices, SynWin required only an off-the-shelf personal computer. The tasks in SynWin were selected to provide a generic work environment where the operator is required to remember and classify items on demand, perform a self-paced task (arithmetic problems), and monitor and react to both visual and auditory information. The result is a prototypical, PC-based synthetic work task. No attempt was made to simulate any particular job or system, although the program provides a reasonable part-simulation of various watch-standing jobs.

Task training on SynWork was conducted during the three days immediately preceding the first Schedule. The participants completed SynWork four times during training, once during

the afternoon of the first training day, once during the afternoon of the second training day and twice during the morning of the final training half-day. Additionally, they completed a refresher SynWork session once during the afternoon of the stabilization period that preceded each of the three 6-day watchstanding periods. Learning of this complex task continued throughout the study.

The Submarine Skills training Network (SubSkillsNet) was a family of independent training simulations distributed by Submarine On Board Training (SOBT) of the Naval Air Warfare Center-Training Systems Division. The package provided simulations of several workstations on a submarine that were designed to enable seamless, networked use for individual or team training. Originally conceived to address collision avoidance training, the growing set of trainers had resulted in an integrated system that could meet a variety of training objectives.

The SubSkillsNet task used here was a modification of the trainer: a research version that enabled the capture of events and subject inputs in spreadsheet format. We used the Submarine Periscope Observation Trainer (SPOT) subtask as our primary research tool. The SPOT was one of six possible SubSkillsNet training environments, providing the user with a functional periscope. The participants were required to call the ranges and angles on the bow for numerous sea contacts. They were also required to identify all contacts, air and sea, with the assistance of an on-screen database viewer.

Scenarios written by the investigators were altered across testing sessions by rotating each scenario 25 degrees. Additionally, contacts were changed mildly in terms of identity. For example, an SSN was changed to an SSBN to keep the contacts novel and to minimize unwanted practice effects. The total number of contacts and angle on the bow was held constant across sessions.

Task training on SubSkillsNet was conducted during the three days immediately preceding the first Schedule. The participants completed SubSkillsNet five times during training, twice during the morning of the second training day, once during the afternoon of the second training day and twice during the morning of the final training half-day. Additionally, they completed a refresher SubSkillsNet session once during the afternoon of the noon-to-noon stabilization period that preceded each of the three 6-day watchstanding periods. Learning of this complex task continued throughout the study.

This was the seminal use of SubSkillsNet as a research instrument. Thus, our primary focus was on implementation, understanding training needs and developing scenarios. Our secondary focus was on the acquisition of fatigue-related data.

Watch, Drill and Training Measurement Sessions

The incorporation of these three kinds of measurement sessions into the overall watchstanding schedules is shown in Appendix B.

Watch Sessions. In each hour of a watch period, each participant would complete the ANAM battery, complete a 30-minute session on SynWork or SubSkillsNet. Thus, all performance

tests were conducted on a 2-h cycle in which there were two repetitions of the ANAM battery and one repetition each of SynWork and SubSkillsNet. In some hours, participants provided oral temperature, blood pressure, heart rate and oxyhemoglobin saturation measures. Except for oral temperature, these vital sign measures were not analyzed.

Drill Sessions. These sessions were inserted to replicate non-watch work time spent in ship's drills. Six hours of drill occurred in each 72-h Period. In each half-hour of a drill session, each participant would complete three repetitions of the simple response time task and one repetition of the running memory task, they would complete a round of physical tasks, as shown in Table 1, and complete the grip strength test. The 50%-levels of effort in Table 1 were determined with respect to existing Navy physical performance standards, as shown in Table 2.

Table 1. The number of exercise repetitions required in each half-hour of drill periods.

Age	Sit-ups	Push-ups
18-29	23	19
30-39	20	16

Table 2. The minimum Navy requirements for exercise repetitions, from OPNAV Instruction 6110.1F, *Navy Physical Readiness Program*.

Age	Sit-ups	Push-ups
18-29	46	37
30-39	40	31

Training Sessions. These sessions were inserted to replicate non-watch work time spent in training sessions. Six hours of training occurred in each 72-h Period. In each half-hour of a training session, each participant would complete one repetition each of the simple response time task, the logical reasoning task and the PVT.

RESULTS AND DISCUSSION

PARTICIPANT CHARACTERISTICS

The general demographics of the nine male submariner participants are shown in Table 3. Two of the nine participants came from fast attack submarines (SSN), while the other seven came from ballistic missile submarines (SSBN). The technical ratings of the participants were: Electrician's Mate, 1; Electronics Technician, 2; Machinist Mate, 4; Sonar Tech, 1; Yeoman, 1.

Table 3. Demographics of the nine male submariner participants.

Parameter	Mean and SD	Range
Age (years)	27 +/- 7	21- 40
Height (cm)	174.3 +/- 4.9	165 - 180
Weight (kg)	79.6 +/- 16.2	59 - 114
Caffeine (drinks/day)	2.4 +/- 1.7	0 - 5.5
Alcohol (drinks/week	4.1 +/- 3.1	0 - 9
Nicotine	1 smoker/1 pack per day 1chewer/no frequency given	
Education level (years)	12.3 +/- 0.7	12 - 14
Navy experience (years)	7.6 +/- 6.8	2 - 21

By all measures, this group appeared to be a collection of normal, enlisted Navy submariners, aged 21 to 40 yr. None appeared to have obvious clinical problems with depression, anxiety, insomnia or excessive daytime sleepiness. None were extreme "owls" or "larks."

Sleep Behavior Characteristics

Reported Sleep Length. The participants reported ideal sleep times of 6 to 8 h (6.8 +/- 0.87 h), but reported that their usual, 24-h sleep totals ranged from 3 to 10 h (6.7 +/- 1.9 h, all in a single period). Thus, it appeared that they usually acquired their ideal amount of sleep, but occasionally experienced acute fatigue from shortened sleep periods. Their reported sleep time was approximately normal. In a National Sleep Foundation poll taken in the year 2000, those surveyed reported sleeping about seven hours a night on the average. About one-third surveyed tended to sleep eight or more hours, and one-third tended to sleep 6.5 hours or fewer. About ¾ of our participants reported taking occasional naps of 20 to 90-min length (56 +/- 27 min).

Reported Sleep Latency. The participants reported usual sleep latencies of 5 to 45 min (19.4 +/- 13.8 min). On a scale of 1 to 7 (1, "Not at all", to 7, "Very much"), they rated their difficulty falling asleep as ranging from 2 to 7 (3.0 +/- 1.7). These numbers suggested that the group did not usually suffer from excessive sleepiness: they did not fall asleep very quickly nor unusually easily. Very sleepy individuals tend to fall asleep in about 5 min, according to the general results of the Multiple Sleep Latency Test used in sleep clinics. Two of the participants reported latencies shorter than 7.5 min, and three reported latencies greater

than 20 min. Six participants reported difficulties less than 3, and one reported difficulty greater than 5.

Reported Sleep Inertia. The participants reported usual sleep inertia lengths of 1 to 20 min (10.7 +/- 6.6 min). On the 1-to-7 scale, they rated their difficulty getting up as ranging from 2 to 7 (3.3 +/- 1.6). These numbers suggested that the group did not usually suffer from excessive inertia: they passed through sleep inertia at an expected rate and did not report great difficulty doing so. One participant reported a sleep inertia less than 5 min, and two reported inertias lasting longer than10 min. Three participants reported difficulties less than 3, and one reported difficulty greater than 5.

Epworth Sleepiness Scale (ESS)

The participants reported ESS ratings ranging from 4 to 12 on the 0-to-24 sleepiness scale (mean 8.0 +/- 2.9; median 7). A higher score represents greater sleepiness. None reported sleepiness above 15, providing no cause for concern with respect to acceptable individual job performance.

Morningness-Eveningness Questionnaire (MEQ)

Scores on the MEQ were used to categorize individuals as follows (Horne and Östberg, 1976):

- Definitely morning, 70-86
- Moderately morning, 59-69
- Neither, 42-58
- Moderately evening, 31-41
- Definitely evening, 16-30

The participants, individually and as a group, exhibited central tendency with a mean score of 50.0 +/- 6.9 and a median score of 51. The scores ranged from 36 to 58, distributed as:

- Neither, 8 participants
- Moderately evening, 1 participant

Sleep Hygiene and Practices (SHAPS)

The participants reported:

- Sleep Hygiene Knowledge ratings ranging from 22 to 31 on the 13-to-39 scale (mean 26 +/- 3.5; median 27).
- Caffeine Knowledge ratings ranging from 47.1 to 83.3 on the 0-to-100 scale (mean 67.4 +/- 12.7; median 68.8).
- Sleep Hygiene Practice ratings ranging from 39 to 53 on the 0-to-133 scale (mean 49.0 +/- 5.2; median 68.8).

Higher scores indicated more knowledge or less healthy sleep hygiene practices. The participants indicated moderately good sleep hygiene knowledge, good caffeine knowledge, and good sleep practices.

Beck Depression Inventory (BDI)

The participants reported BDI ratings ranging from 1 to 14 on the 21-point depression scale (mean 5.4 +/- 4.6; median 4). A higher score represents greater depression. The group fell in the 4 to 7 range, defined as normal or as mildly depressed. Three participants scored in the 8 to 15 range, defined as approximately moderately depressed

Trait Anxiety Inventory (TAI)

The participants reported TAI ratings ranging from 26 to 45 on the 20-to-80-point anxiety scale (mean 35.3 +/- 5.9; median 36). A higher score represents greater anxiety. The trait anxiety norm for working men, aged 19 to 39 yr, is 35.55 +/- 9.76. The group data replicated this norm.

Cognitive Hardiness Scale (CH)

The participants reported CH ratings ranging from 98 to 122 on the 30-to-150-point hardiness scale (mean 108.4 +/- 9.3; median 108). A higher score represents greater cognitive hardiness. The group mean and median fell near the population mean of about 106 and very slightly lower than a military mean of about 114 to 118.

Correlations

The intercorrelation matrix (Spearman r) for self-reported ideal sleep length, self-reported sleep latency length, self-reported sleep inertia length, BDI score, TAI score, and ESS score is shown in Table 4. The only statistically significant relationship occurred between the TAI and the ESS (Spearman $r = 0.710$, $p = 0.032$), suggesting that the two measures (trait anxiety and general sleepiness, respectively) shared about 50% variance. Generally, then, this subset of intake measures overlapped minimally in the kinds of information they elicited from the participants.

Table 4. Intercorrelation matrix (Spearman r) for reported ideal sleep length, reported sleep latency length, reported sleep inertia length, BDI score, TAI score, and ESS score ($n = 9$; $*p < 0.05$).

	Length	**Latency**	**Inertia**	**BDI**	**TAI**
Latency	-0.424				
Inertia	-0.215	0.035			
BDI	-0.072	0.286	-0.238		
TAI	0.009	0.030	-0.090	0.424	
ESS	0.382	0.155	0.083	0.211	0.710*

PERFORMANCE TESTING

The first 72-h, second 72-h, and recovery data collection periods were labeled P1, P2, and R, respectively. As with all extended studies, learning occurred continuously in many tasks, even where least expected. Examinations of time-series plots showed that the learning occurred at a slow, steady pace. We dealt with this problem by using pre-Schedule baselines to adjust for the major part of the learning that occurred. A baseline was established during the first watch of P1. The baselines differed significantly (paired t test, df = 8, $p < 0.05$) across Schedules for several measures (Table 5).

Table 5. Significant differences in performance task baselines (paired t test, df = 8, $p < 0.05$) across Schedules. Pairs include the Alternative (A), Maritime (M) and Submarine (S) watch schedules. RT is response time. SD is standard deviation. FRRT is the reciprocal of the fastest RTs.

Test	Measure	Pair	Mean Difference	2-Tail t	p
Logical	Accuracy	M-S	-1.88%	-2.47	0.039
Logical	SD RT, correct	A-M	-170 ms	-2.48	0.038
Match to Sample	Accuracy	A-M	4.92%	2.54	0.034
Match to Sample	Accuracy	A-S	5.96%	2.88	0.020
Match to Sample	SD RT, correct	A-M	-198 ms	-2.38	0.045
Mental Arithmetic	Mean RT, correct	A-S	438 ms	2.37	0.045
Mental Arithmetic	Mean RT, correct	M-S	173 ms	2.48	0.038
Mental Arithmetic	No. of omissions	A-S	0.48	2.42	0.042
Mental Arithmetic	Throughput	A-S	-9.46/min	-2.39	0.044
PVT	Mn_FRRT	A-S	0.51/msec	2.72	0.026
PVT	Mn_RRT	A-M	0.62/msec	2.50	0.037
PVT	Mn_RRT	A-S	0.53/msec	2.59	0.032
PVT	Mn_SRRT	A-M	0.63/msec	2.47	0.039
PVT	No. of lapses	A-M	-5.6	-3.46	0.009
Simple RT	Throughput	A-S	-16.4/min	-3.26	0.046

Note that in almost all cases, alert, baseline performance improved across Schedules, with respect to time. For example, throughput was greater in S than in A for the Simple RT task, mean RT was shorter in S than in A for Mental Arithmetic, etc. Exception: the number of baseline lapses on the PVT was greater in the M Schedule than in the A Schedule. Obviously, even on these reasonably simple tasks, some learning occurred across the month of testing.

The performance data were reported as differences from the watch no. 1 baseline performance level (P1 minus baseline, P2 minus baseline, and R minus baseline; within subjects). Thus, positive differences indicated increase from baseline. A mean was computed for all change data from within watches for P1 and P2 (P1 did not include the baseline watch period) and for the recovery day. Thus, the analytical method was a 2-factor, 3- x 3-level (Schedules and Periods) analysis of variance (ANOVA) with repeated measures on both factors and without randomization of orders of presentation. The three Schedules, in order of presentation, were Alternate (A), Maritime (M) and Submarine (S). The three periods, in

order of presentation, were P1, P2 and R. All tests were corrected for sphericity errors by the Huyn-Feldt procedure. Significant main effects of Schedules and the interactive effects of Schedule by Period are reported here ($p < 0.05$)[5]. The main effects of Periods were of low interest for this experiment; thus, the slight imbalance in numbers of combined watches between P1 and P2 was not a concern in this analysis. *Post hoc* analyses were conducted using the Newman-Keuls procedure.

One advantage of this method of summarizing the data was that performance during all possible watches in one schedule was compared with performance during all possible watches in another schedule. Note, however, that only the 12-to-4 watch period was assessed during the Maritime (M) Schedule.

Simple Response Time

There were no significant effects on changes in accuracy, mean response time, omissions, the standard deviation of response time, or throughput.

Mental Arithmetic

There were no significant effects on changes in accuracy, mean response time, omissions, the standard deviation of response time, or throughput.

Delayed Matching-To-Sample

There were no significant effects on changes in accuracy, mean response time, omissions, the standard deviation of response time, or throughput. The main effect of Schedule on change in SDRTC approached significance (H-F corrected $p = 0.053$). Post hoc tests indicated Schedule A values were greater than Schedule S values. Thus, the variability of response time appeared to be greater during the first session (A) than during the last session (S), several weeks later. This may have been due to the different watchstanding schedules or, more likely, due to a learning effect.

Logical Reasoning

There were no significant effects on changes in accuracy, mean response time, omissions, the standard deviation of response time, or throughput.

Running Memory

There were no significant effects on changes in accuracy, mean response time, the standard deviation of response time, or throughput. There was a significant interactive effect of Schedule and Period on the change in number of omissions ($F(4,32) = 3.72$, H-F MSe =

---

[5] In each description of a significant effect throughout this report, the raw degrees of freedom (df) are listed instead of the Huyn-Feldt (H-F)-corrected degrees of freedom (e.g., F(2,16) for 2 and 16 raw df). This was done to allow the reader to understand more clearly the structure of the ANOVA. However, the H-F-corrected df, error term values and p values were used for all assessments of significance and *post hoc* tests..

0.751, H-F p = 0.044). There was a significant main effect of Periods (F(2,16) = 7.71, H-F MSe = 1.615, H-F p = 0.018) but no significant main effect of Schedule (F(2,16) = 0.371, H-F MSe = 5.827, H-F p = 0.662). The *post hoc* analysis indicated that the increase from baseline level in the number of omissions was significantly greater during P1 of the S Schedule than the decreases that occurred during all three post-recovery-sleep periods (S_P1 >> A_R, M_R and S_R; $p < 0.05$; Figure 4). The gross difference across the Schedule S, Period P1, change and the Schedule A, Period R, change was about two omissions on a task that presented about 40 stimuli (about a 5% gross difference). This result suggested that the participants experienced their poorest performance on the Running Memory Task (at least, in terms of omissions) during P1 of the familiar S schedule, and did experience some performance recovery as a result of the recovery period.

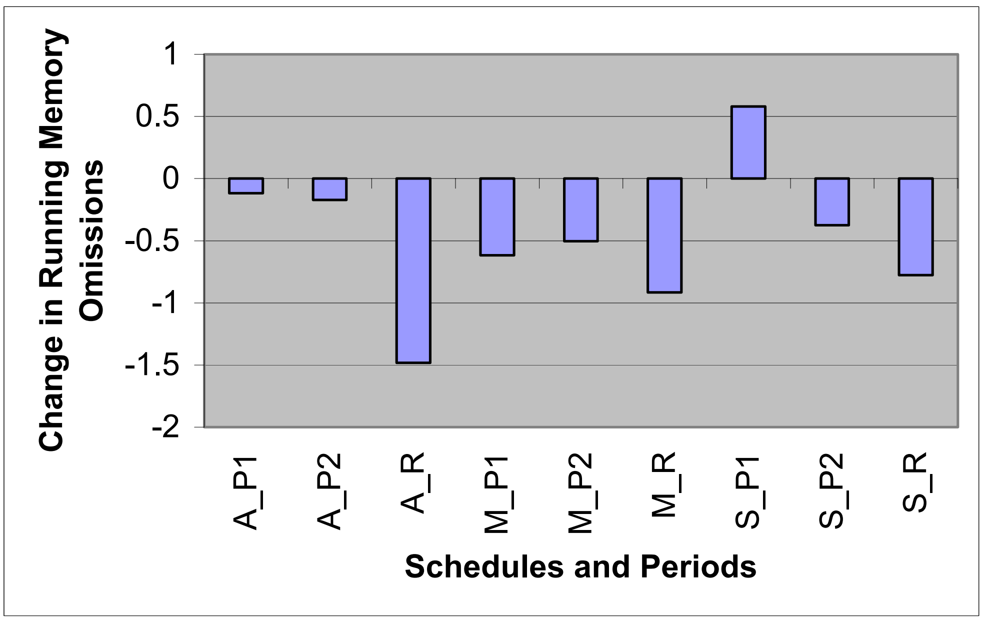

Figure 4. Interactive effects of Schedules (A, M, S) and Periods (P1, P2, R) on omissions during the Running Memory Task.

Psychomotor Vigilance Task

There were no significant effects on changes in lapses, Mn_FRRT, Mn_SRRT, or SDRT. There was a significant main effect of Schedule on change in Mn_RRT (F(2,16) = 3.89, H-F MSe = 0.291, H-F p = 0.042), but the effect of the interaction between Schedule and Period was not significant. The *post hoc* analysis indicated that the increase in Mn_RRT from baseline during the M Schedule was significantly different than the decrease associated with the A Schedule (M >> A; $p < 0.05$; Figure 5). The reciprocal RT increased from baseline during the M Schedule, while it decreased from baseline during the A Schedule. Thus, conversely, the RT decreased from baseline during the M Schedule, while it increased from baseline during the A Schedule.

The total difference in Mn_RRT change between the M and A Schedules shown in Figure 5 represents a 26-msec (approximately 10%) difference, based upon a grand mean baseline RT

of 269 msec[6]. Thus, it appeared that the M schedule provided a small advantage in response time on the PVT.

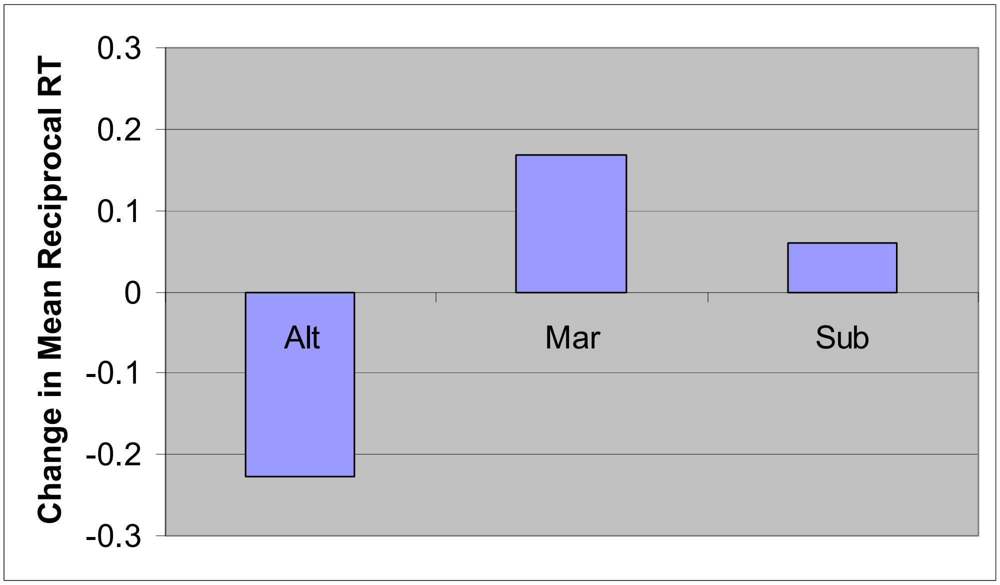

Figure 5. Schedule effect on the change in the mean reciprocal response time (Mn_RRT) for the Psychomotor Vigilance Task.

SynWork

Only the composite SynWork score was assessed for changes. A baseline was established during the first watch of P1. Due to a continuous, monotonic, nearly linear learning process that continued throughout the experiment, the baselines differed significantly from the A to the M Schedules (2-tail, paired $t(8) = -3.00$, $p = 0.017$) and from the M to the S Sessions ($t(8) = -3.84$, $p = 0.005$). The mean Schedule scores were 6430.8 +/- sd 1077.3, 7641.8 +/- 1652.2, and 8169.3 +/- 2131.3 for Schedules A, M and S, respectively.

The SynWork data were also reported as differences from the Schedule's watch no. 1 baseline performance. As for the ANAM tasks, a mean was computed for all data from within watches for P1 and P2 (P1 did not include the baseline watch period) and for the recovery day. Thus, the analytical method was a 2-factor, 3- x 3-level (Schedules and Periods) analysis of variance (ANOVA) with repeated measures on both factors and without randomization of orders of presentation. All tests were corrected for sphericity errors by the Huyn-Feldt procedure. Significant main effects of Schedules and the interactive effects of Schedule by Period were sought. There were no significant main effects of Schedule nor interactive effects of Schedule by Period on changes in the SynWork composite score.

SubSkillsNet

[6] Note that, because of the nonlinear relationship between the measurement domain of the reciprocal transform and the time domain, the translation of effect size was not straightforward. In fact, the difference was 9.7% of baseline in the time domain, and 10.6% in the reciprocal domain.

Because this was a seminal effort using this package, the first schedule (A) was used to determine proper difficulty level of SPOT periscope task scenarios and the usability of the SurfCAT task. We determined that SurfCAT, a driving task involving maneuvering a submarine through various simulated channels, would not be a good experimental measure due to the excessive time required to realistically traverse a waterway. Therefore, this analysis involved a comparison of SPOT data only between the latter two schedules. There were no significant effects on contacts found or on accuracy in estimating angle on the bow between schedules M and S.

SUBJECTIVE MEASURES

These data were also reported as intra-subject differences from baseline performance. The data from the P1, P2 and R Periods were compared to the data from the first-watch baseline, as with the performance data, or to the baseline established during the Stabilization day. For state anxiety and mood, a pre-watch, post-watch factor was added to the analysis. For symptoms, the non-parametric Friedman two-way analysis of variance was used.

Stanford Sleepiness Scale

There were no significant effects of Schedule or Schedule by Period interaction on changes in SSS ratings. There was a significant effect of Period on the change in the SSS rating ($F(2,16)$, H-F MSe = 0.238, H-F $p = 0.000$). A *post hoc* analysis indicated that the decline measured in the R Period was significantly greater than the declines measured in P1 and P2 (R >> P1 and P2; $p < 0.01$; Figure 6). After recovery sleep, the mean SSS rating declined about one scale unit from baseline on the 1-to-7 SSS scale. The mean baseline rating was 2.9. Thus, the participants improved monotonically from a mean baseline rating of about three ("Relaxed; awake; not at full alertness; responsive") to about two ("Functioning at a high level, but not at peak; able to concentrate") after recovery sleep. This result suggested that the participants felt slightly, but reliably, less sleepy after recovery sleep than during P1 and P2. This finding suggests that the SSS was sensitive to at least one fixed factor during this experiment.

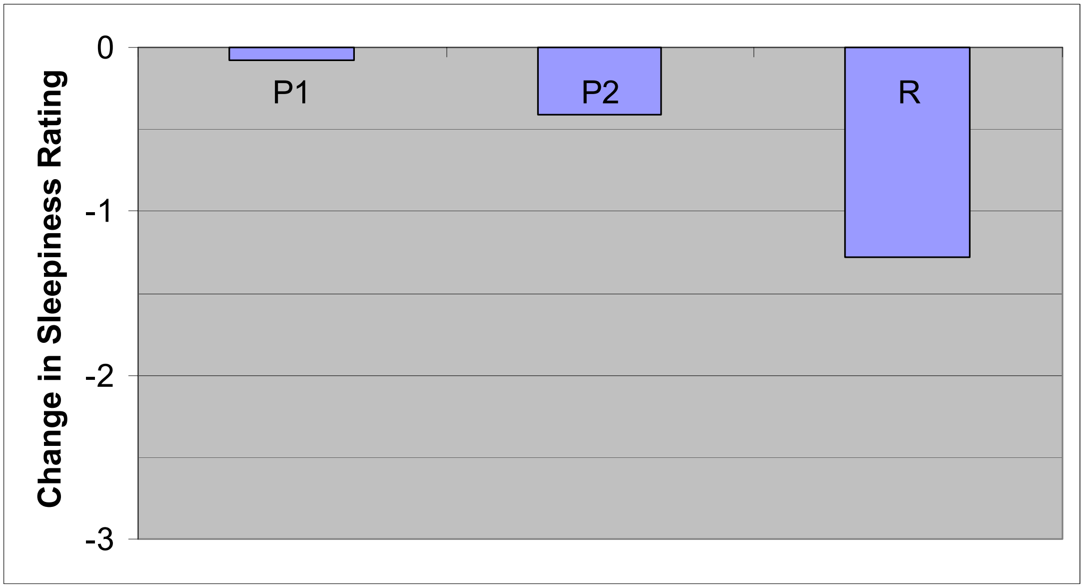

Figure 6. Period effect on the change in the Stanford Sleepiness Scale rating.

Mood

Due to missing data, separate 2-factor ANOVAS (Schedule, Pre-Post-Watch) were run for the B ($n = 6$), P1 ($n = 9$), P2 ($n = 9$), and R ($n = 6$) Periods. Main effects for these two factors and the interactive effects are reported here, with significance accepted at $p < 0.05$. For the B Period, there were no significant main effects of Schedule nor significant interactions. However:

- There was a significant main Pre-Post effect on Activity ($F(1,5) = 11.03$, H-F MSe = 157.4, H-F $p = 0.021$). The participants reported being less active after the watch.
- There was a significant main Pre-Post effect on Fatigue ($F(1,5) = 19.29$, H-F MSe = 202.5, H-F $p = 0.007$). The participants reported being more fatigued after the watch. This was a classic pattern of acute fatigue.
- There was a significant main Pre-Post effect on Happiness ($F(1,5) = 6.79$, H-F MSe = 137.4, H-F $p = 0.048$). The participants reported being less happy after the watch.

For P1, there were no significant main effects of Schedule nor significant interactions. However, again:

- There was a significant main Pre-Post effect on Activity ($F(1,8) = 77.53$, H-F MSe = 35.53, H-F $p = 0.000$). The participants reported being less active after the watch.
- There was a significant main Pre-Post effect on Fatigue ($F(1,8) = 55.34$, H-F MSe = 57.95, H-F $p = 0.000$). The participants reported being more fatigued after the watch. This was a classic pattern of acute fatigue.
- There was a significant main Pre-Post effect on Happiness ($F(1,8) = 23.16$, H-F MSe = 74.09, H-F $p = 0.000$). The participants reported being less happy after the watch.

For P2, the pattern changed:

- There was a significant interactive effect on Activity ($F(2,16) = 5.06$, H-F MSe = 102.54, H-F $p = 0.032$) and a significant main effect of Pre-Post-Watch (Figure 7).

The *post hoc* test indicated that pre-watch Activity was reported to be highest before watches in the A and S schedules, and lowest after watches in the A Schedule (A_Pre, S_Pre >> all Post and M_Pre; S_Pre >> A_Pre; A_Post << M_Post, S_Post; $p < 0.05$). The latter effect on A_Post was likely due to the compression of 12 h of work into 18 h in the A Schedule.

- Similarly, there was a significant interactive effect on Fatigue ($F(2,16) = 4.45$, H-F MSe = 128.34, H-F $p = 0.029$) and significant main effects of Schedule and Pre-Post-Watch (Figure 8). The *post hoc* test indicated that Fatigue was reported to be highest in the M Schedule (pre and post) and lowest in the A Schedule (pre and post) (M_Post >> all; M_Pre, S_Pre >> A_Pre, A_Post; M_Pre >> S_Pre; $p < 0.05$). These effects may have represented the malaise associated with circadian rhythm disorder expected in the M Schedule, contrasted with the expected circadian stability of the A Schedule.

For R, there were no significant main effects of Schedule nor Pre-Post-Watch nor significant interactions. The lack of a Pre-Post effect was likely due to the participants' anticipation of leaving the confines of the laboratory after 8 d.

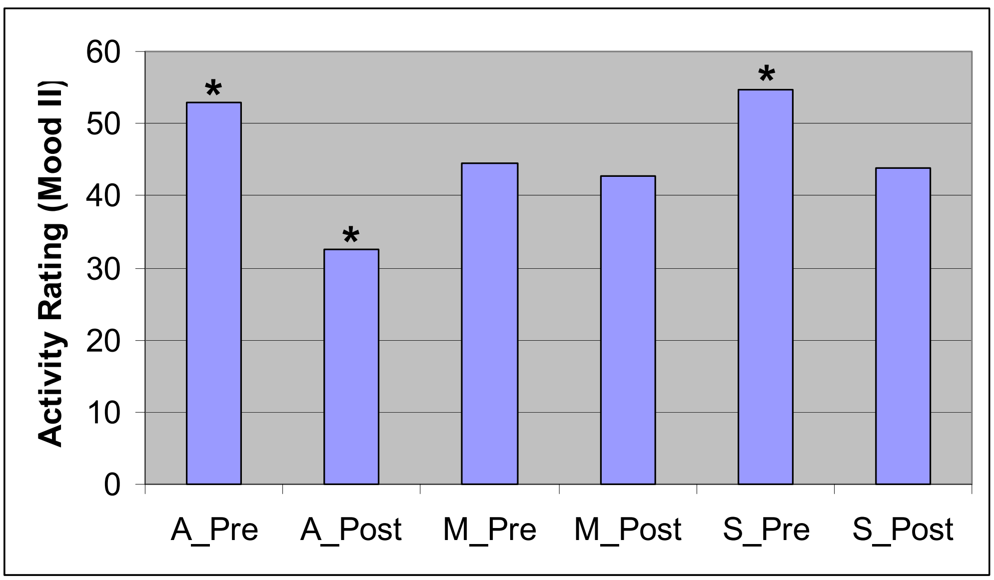

Figure 7. Significant interactive effects of Schedule and Pre-Post-Watch on Activity reported in the Mood 2-R scale (A_Pre, S_Pre >> all Post and M_Pre; S_Pre >> A_Pre; A_Post << M_Post, S_Post; *$p < 0.05$). Schedules are A, M and S. Pre and Post refer to before and after watches.

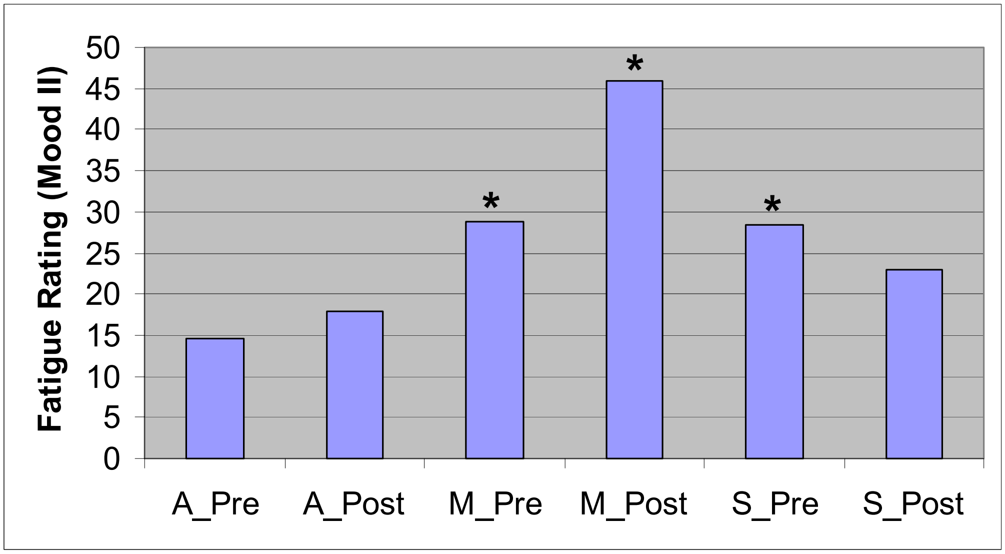

Figure 8. Significant interactive effects of Schedule and Pre-Post-Watch on Fatigue reported in the Mood 2-R scale (M_Post >> all; M_Pre, S_Pre >> A_Pre, A_Post; M_Pre >> S_Pre; *p < 0.05). Schedules are A, M and S. Pre and Post refer to before and after watches.

Symptoms

Symptom reports were reduced to the mean value of all reports (ratings of 1 through 7) by an individual within a Period. Most often, this value was zero. The means were reduced to ranks and the non-parametric Friedman two-way analysis of variance was used to compare Schedules within each of the four Periods (B, P1, P2, and R; n = 9 for all Periods). Significance was accepted at $p < 0.05$. There were no significant effects for the B Period. During P1, the effect on "Irritability" was significant ($Chi^2(2) = 7.54$, $p = 0.023$), and the effect on "Vivid dreams" reached the 0.05 level ($Chi^2(2) = 6.00$, $p = 0.050$). In both cases, the highest rank occurred during the A Schedule.

During P2, the effect on "Trouble staying awake" was significant ($Chi^2(2) = 8.00$, $p = 0.018$), as was the effect on "'Drugged' feeling" ($Chi^2(2) = 7.43$, $p = 0.024$). Again, in both cases, the highest rank occurred during the A Schedule. There were no significant effects for the R Period.

Pre-Sleep Arousal

Raw scores, not change scores, were assessed for the cognitive and somatic scales. The 2-factor ANOVA was run on the baseline (Stabilization day), P1 and P2 data. Due to missing data, the Recovery Period was handled with a paired t test that compared the A and S sessions. There were no significant main effects of Schedule or interactive effects on the cognitive nor the somatic scores on the Pre-Sleep Arousal Survey. The somatic pre-sleep score for Recovery after the A Schedule was significantly greater than the somatic pre-sleep score for Recovery after the S Schedule ($T(8) = 2.82$, $p = 0.023$; 7.3 score >> 6.1 score).

There was a significant main effect of Period on the somatic score. By itself, this effect was not relevant in the context of this report. However, the plot of the Schedule by Period interaction was of great relevance, even though the interactive effects did not achieve statistical significance (Figure 9). The highest pre-sleep somatic score values were reported for the second 72-h Period of the M Schedule. The terrorist acts of September 11, 2002, occurred less than 24 h after the start of this Period, and were reported immediately to the study participants, as described above.

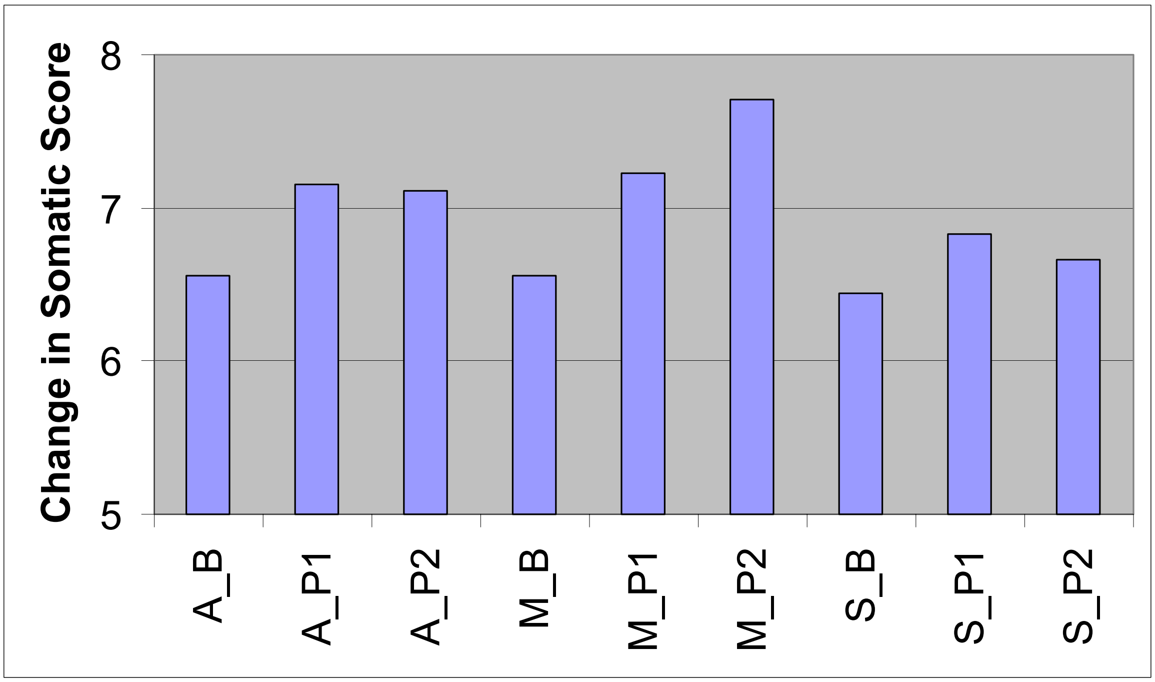

Figure 9. Interactive effects of Schedules (A, M, S) and Periods (B, P1, P2) on the somatic score of the Pre-Sleep Arousal Survey.

State Anxiety

The main effects of Schedule and Pre-Post-Watch, the two-way interactions and the three-way interaction were not statistically significant for the state anxiety score. There was a significant main effect of Period that was irrelevant in the context of this report.

Mental Workload Ratings, Physical Workload Ratings

Generally, perceptions of both mental and physical workload declined significantly with time across the three conditions. This was true for the baseline day, P1, P2, and the recovery day. The perceptions were usually higher in the first session (Alternate schedule), relatively lower in the second session (Maritime) and lowest in the third session (Submarine). At no time was the average absolute perception of either kind of workload high. Mean mental workload perceptions ranged from 3 to 4 on the 7-point scale, and mean physical workload perceptions ranged from about 1.5 to 2 on the 15-point scale. Since the effect of sessions cannot be separated from the effect of time and training, little further can be said about these data.

PHYSIOLOGICAL MEASURES

Polysomnography (PSG)

The analysis of the PSG data was hampered slightly by missing data. Technical problems with our Medilog ambulatory recorders[7] caused the loss of some data. Thus, slightly modified, single-factor ANOVA structures were used, as described here.

Stabilization Night. During the 24-h stabilization period (noon to noon), the participants spent their first night in the CASL at the start of Schedule A, and then similar first nights for Schedules M and S. These data are presented descriptively, not analytically, since the participants were adjusting to the new sleep environment to one degree or another (Table 6). The TST for these Stabilization sleep periods were slightly above the expected level and their SE and %WASO were greater than expected for men in this age range (Williams *et al.*, 1974). A single-factor, 3-level ANOVA with repeated measures and Huyn-Feldt correction indicated no significant differences across Schedules for any of the Stabilization sleep measures.

Table 6. Characteristics of sleep during the Stabilization period preceding Sessions A, M and S. SLat is sleep latency, SE is sleep efficiency, TST is total sleep time, WASO is wake time after sleep onset, SREM is rapid eye-movement sleep.

Measure	Alternate	Maritime	Submarine
SLat (min)	18.71 +/- 7.31	29.50 +/- 32.00	23.31 +/- 17.64
SE (%)	94.79 +/- 2.83	87.99 +/- 11.83	89.08 +/- 10.48
TST (h)	7.58 +/- 0.23	7.04 +/- 0.95	7.13 +/- 0.84
%WASO	3.39 +/- 2.60	6.74 +/- 8.08	3.46 +/- 3.46
%S1	3.32 +/- 0.97	4.82 +/- 1.20	5.82 +/- 3.78
%S2	62.84 +/- 11.43	65.48 +/- 9.78	59.18 +/- 6.08
%S3	4.68 +/- 2.76	3.72 +/- 3.15	3.95 +/- 3.50
%S4	8.78 +/- 8.19	5.90 +/- 7.55	8.05 +/- 7.35
%SWS	14.36 +/- 10.48	8.06 +/- 10.55	11.39 +/- 7.74
%SREM	20.38 +/- 6.75	20.06 +/- 6.65	23.00 +/- 4.93

Period P1. Data from eight participants were available. These were subjected to a single-factor, 3-level (Schedule) ANOVA with repeated measures. All tests were corrected for sphericity errors by the Huyn-Feldt procedure. Significant main effects of Schedules are reported here ($p < 0.05$). There were significant effects of Schedule on six of nine PSG measures assessed. These effects and the results of the respective *post hoc* tests are summarized in Table 7 and Figures 10 through 12. There were no significant effects of Schedule on percent Stage 1 (%S1), percent Stage 3 (%S3), or sleep latency (SLat).

[7] Now replaced.

Table 7. Significant effects of Schedule on polysomnography measures during P1 (df = 2, 14). H-F is Huyn-Feldt; WASO is Wake after Sleep Onset); A is Alternate, M is Maritime and S is Submarine Schedule.

Measure	F	H-F MSe	H-F p	*Post hoc*
%S2	8.13	53.157	0.011	A >> M, S; M >> S
%S4	6.13	7.681	0.022	A >> M, S
%SWS	5.62	29.17	0.016	S >> M, A, M >> A
%WASO	4.56	84.508	0.046	M >> S, A
SE	8.17	46.391	0.004	A >> M, S; A >> S
%SREM	6.42	9.677	0.010	S >> M, A
TST	28.37	2.386	0.000	A >> M, S; A >> S

Two effects are shown in Figure 10. First, the total sleep time was significantly lower during the M schedule because of the limited amount of time spent in bed, 21 hours instead of 24 hours per 72 hours, as described, above. To place this effect in context, the average daily amount of sleep obtained in the S Schedule was less than the 7.0 h expected for men of these ages, and the amount in the A Schedule was slightly more (*ibid.*). The lesson here is that, given more time in bed in the A and S schedules, the participants did sleep. This tendency usually suggests the presence of some cumulative sleep debt.

Second, sleep efficiency was also significantly lower during the M schedule (Figure 10). Sleep efficiency for men tends to decline from about 96% to about 91% across the age ranges represented here (Williams *et al.*, 1974). It was relatively low in all three Schedules. It is unlikely that the more extreme sleep disruption in the M Schedule was due to schedule irregularity, since sleep occurred at the same time each day. More likely, it was due to a circadian desynchrony similar to that predicted by the SAFTE model (Figure 2, above).

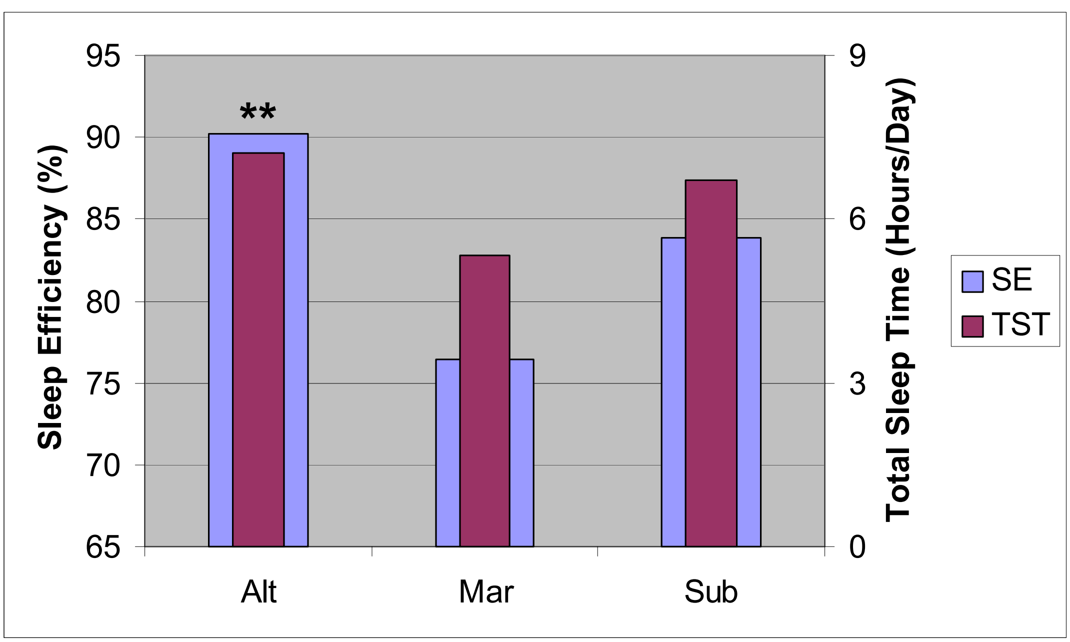

Figure 10. Effect of Schedule on sleep efficiency (SE) and total sleep time (TST) during Period 1 (A >> M, S for both measures; **p < 0.01).

The converse of the lower sleep efficiency during the M schedule is shown in Figure 11. The relative amount of time spent awake after sleep onset (WASO) was significantly higher during the M schedule. This value tends to increase from about 1.5% to 6% in men of these ages (*ibid.*). It was higher than that in all three Schedules, and quite high in the M Schedule.

The relative amount of time spent in stage 2 sleep was slightly high during the A Schedule (Figure 11). Generally, this number should be about 45% to 55% (*ibid.*), as seen in the M and S Schedules.

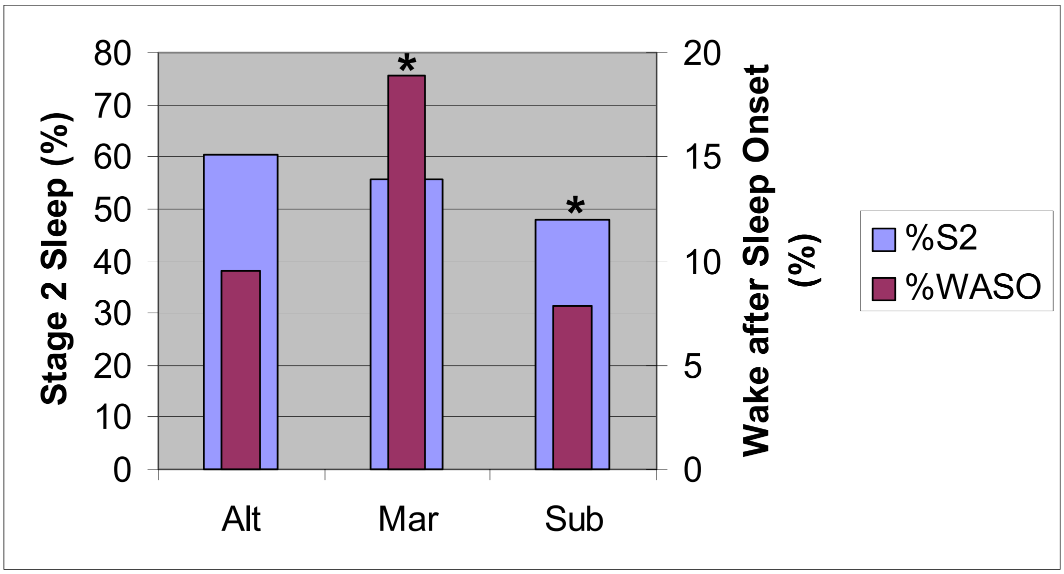

Figure 11. Effect of Schedule on percent stage 2 sleep and wake after sleep onset (WASO) during Period 1 (A >> M, S; M >> S for %S2; M >> S, A for %WASO; *$p < 0.05$).

Generally the proportion of time spent in stage 4 sleep should be about 7 to 14% for men of these ages (*ibid.*) It appeared to be suppressed slightly in the A and M Schedules, as did SWS (Figure 12). The proportion of time spent in stage REM sleep should be about 23 to 28%. Again, it appeared to be suppressed slightly in the A and M Schedules. The higher relative amounts of stages 4 and REM sleep during the S schedule shown in Figure 12 may have been a result of the constricted sleep period (6 h) used throughout the S schedule.

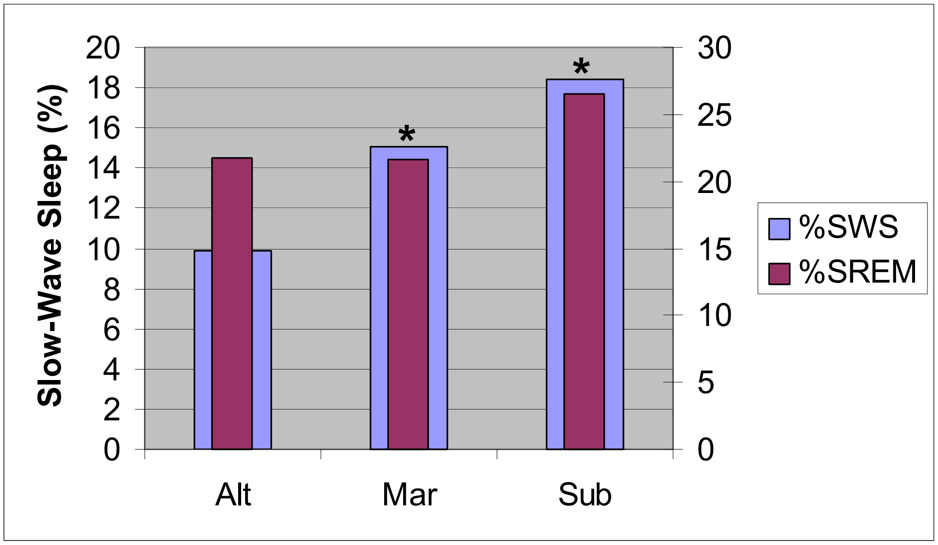

Figure 12. Effect of Schedule on percent SWS and REM sleep during Period 1 (S >> M, A for both measures; *$p < 0.05$).

Period P2. Data from nine participants were available. The same ANOVA was applied. Significant main effects of Sessions are reported here ($p < 0.05$). There was a significant effects of Schedule on only one of the nine PSG measures assessed, TST ($F(2,16) = 16.98$, H-F MSe = 2.092, H-F $p = 0.000$). The *post hoc* tests indicated that TST was significantly greater during Schedule A than during Sessions M and S (A >> M, S; $p < 0.01$). This effect is shown in Figure 13.

The participants acquired 2.7 and 4.0 hours more sleep, across the second 72 hours of the Alternate schedule than in the Maritime and Submarine schedules, respectively. The A Schedule was the only Schedule in which the expected 7.0 h of TST was achieved. There were no significant effects of Schedule on percent SLat, %S1, %S2, %S3, %S4, %WASO, %SREM, or SE, so one may say, although that sleep quality was approximately equal across the three Schedules, more good-quality sleep was acquired in the Alternate schedule.

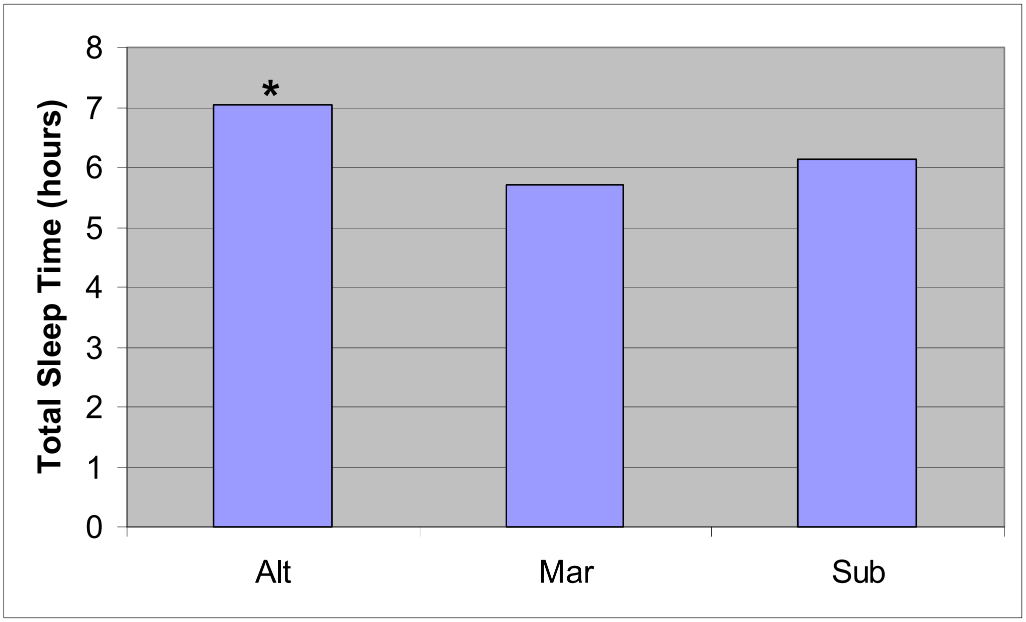

Figure 13. Effect of Schedule on total sleep time during Period 2 (A >> M, S; *$p < 0.05$).

Recovery. Data from six participants were available. The same ANOVA was applied. Significant main effects of Schedules are reported here ($p < 0.05$). There were significant effects of Schedule on %S2 ($F(2,10) = 6.66$, H-F MSe = 74.750, H-F $p = 0.033$), and on %S3 ($F(2,10) = 6.85$, H-F MSe = 3.532, H-F $p = 0.013$). The *post hoc* tests indicated that %S2 was significantly greater during Schedule A than during Schedules M and S (A >> M, S; $p < 0.05$), and that %S3 was significantly greater during Schedule M than during Schedules S and A, and significantly greater during Schedule S than during Schedule A (M >> S, A; S >> A; $p < 0.05$). These effects are shown in Figures 14 and 15. There were no significant effects of Schedule on percent SLat, %S1, %S4, %WASO, %SREM, TST, or sleep efficiency.

The proportion of time spent in stage 2 sleep should be about 45% to 55%, and this amount was achieved in the R Period after participation in all three Schedules (Figure 14). The proportion of time spent in stage 3 sleep should be about 6% (*ibid.*). This proportion was not achieved in any of the R Periods, though the post-M recovery sleep came close (Figure 15) The sleep patterns displayed during Recovery suggested that the Alternate schedule created a lesser need for slow-wave sleep (S3) than the other two schedules, and that stage 2 sleep may have substituted for slow-wave sleep after the Alternate watchstanding schedule.

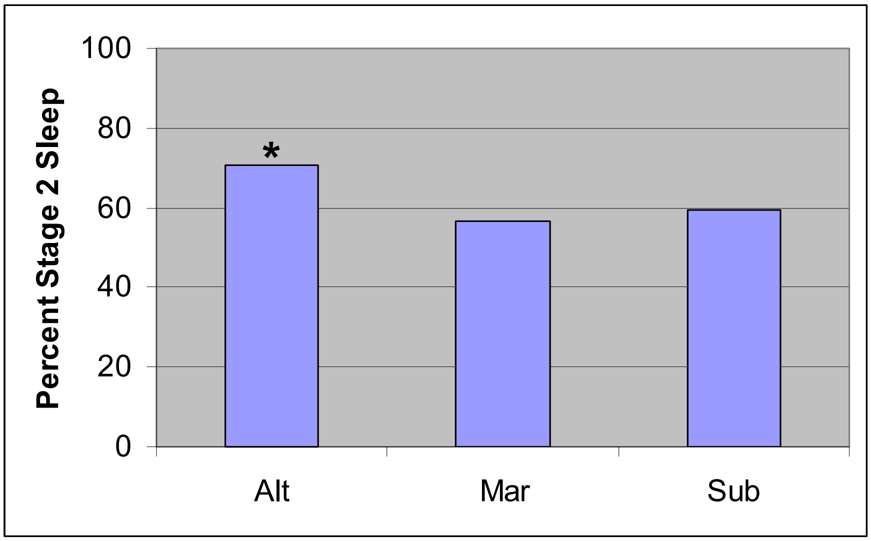

Figure 14. Effect of Schedule on percent Stage 2 sleep during Recovery (Alt >> Mar, Sub; *p < 0.05).

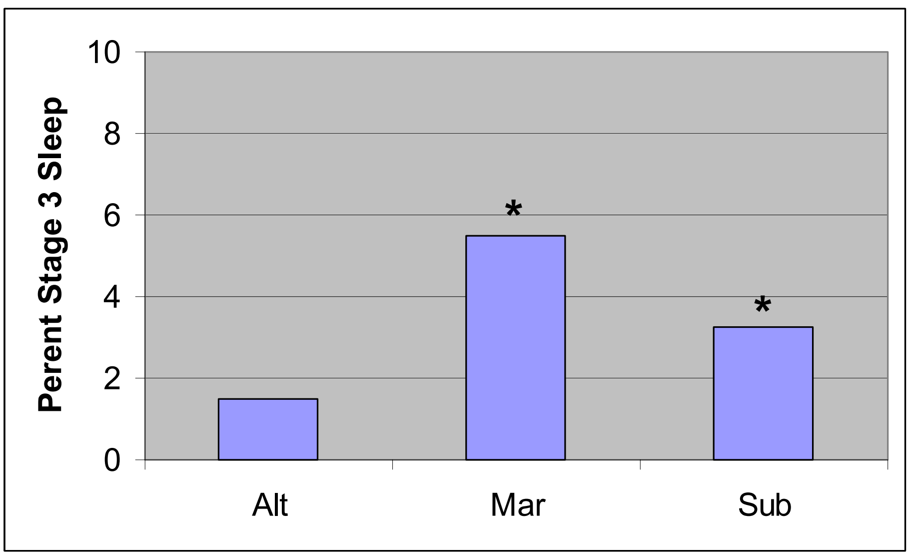

Figure 15. Effect of Schedule on percent Stage 3 sleep during Recovery (Mar >> Alt, Sub; Sub >> Alt; *p < 0.05).

Activity

Data from actigraphy were used to track the participant's sleep patterns in the laboratory and for 72 h before and after each Stabilization and Recovery day, respectively. These were periods during which the participants were living in the local area, outside the laboratory (except after participation in the S schedule). Only the pre- and post-participation data are reported here. They were reduced to the numbers of minutes slept per 72 h and subjected to a single-factor, 3-level (Schedules) ANOVA with repeated measures.

Before participation, the main effect of Schedules was significant (F(2,16) = 4.34, H-F MSe = 89092.8, H-F p = 0.031). The *post hoc* test indicated that significantly less sleep was generated before the A Schedule than before the M and S Schedules (Figure 16; M, S >> A; p < 0.05). This effect was most likely due to the immediacy of the participants' arrival in San Antonio to participate in the 39-day temporary duty assignment to the laboratory.

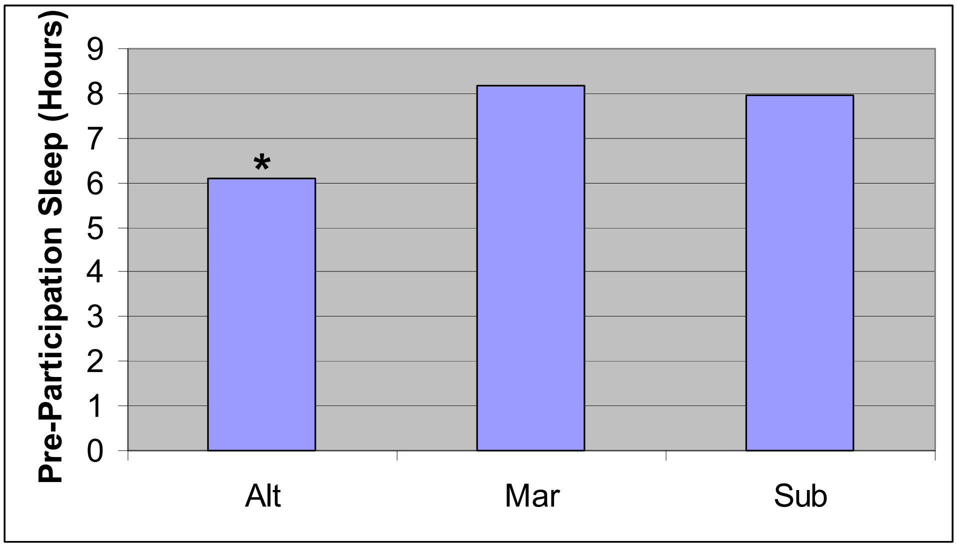

Figure 16. Effect of Schedule on pre-participation sleep, shown in hours per day, averaged across three days (M, S >> A; p < 0.05).

After participation, the main effect of Schedules was significant (F(2,16) = 9.09, H-F MSe = 45442.8, H-F p = 0.003). The *post hoc* test indicated that significantly more sleep was generated after the M Schedule than after the A or S Schedules, and that significantly more sleep was generated after the A Schedule than after the S Schedule (Figure 17; M >> S, A; A >> S; p < 0.01). Note that participants traveled home from the 39-day temporary assignment to the laboratory occurred immediately after their participation in the S Schedule. Discounting the effects of traveling home on the sleep acquired after study participation, it appeared that a greater degree of recovery was required after the M Schedule than after the A Schedule.

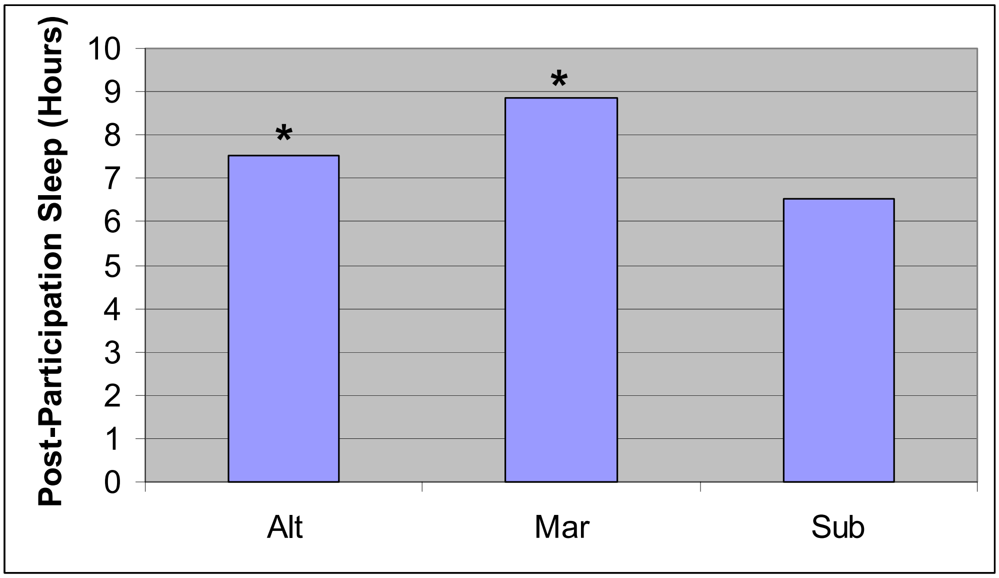

Figure 17. Effect of Schedule on post-participation sleep, shown in hours per day, averaged across three days (M >> S, A; A >> S; $p < 0.01$).

Oculometry

We expected to find cumulative fatigue effects expressed as reductions in baseline pupil size and saccade velocity, and increased pupil response latency. Comparisons were made to the baseline established during the 24-h Stabilization period before each of the three Sessions. There was a significant interactive effect of Schedule by Period on initial pupil size ($F(4,28) = 3.74$, H-F MSe = 0.110, H-F $p = 0.029$). The *post hoc* test indicated that the change from baseline in initial pupil diameter was significantly different during Period 2 than during the Recovery Period of the M Schedule (M_P2 >> M_R; $p < 0.05$) and actually was reversed in polarity from all other conditions. This effect is shown in Figure 18. The significantly larger pupil size associated with Period 2 of the M Schedule was consistent with generalized, elevated sympathetic tone and/or generalized, decreased parasympethic tone. The terrorist acts of September 11, 2002, occurred less than 24 h after the start of this Period, and were reported immediately to the study participants, as described above. There were no significant main effects of Schedule nor interactive effects of Schedule by Period on changes in saccade velocity nor pupil response latency

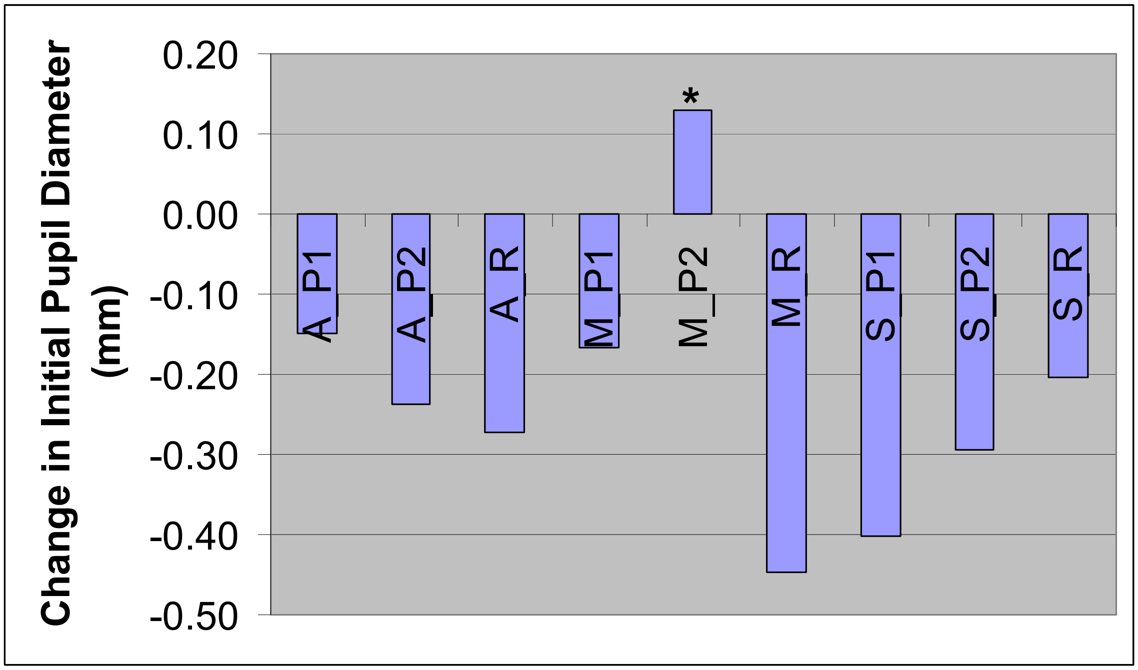

Figure 18. Effects of Schedule and Period on change in initial pupil size (M_P2 >> M_R; $p < 0.05$). Sessions are A, M and S. Periods are P1, P2 and R.

Grip Strength

Force increased monotonically across the three Sessions. The gradual increase was probably due to a combination of both strength and skill development. Comparisons were made to the baseline established during the 24-h Stabilization period. There were no significant main effects of Schedule nor interactive effects of Schedule by Period on changes in the mean or maximum forces generated.

Postural Sway

There were no significant main effects of Schedule nor interactive effects of Schedule by Period on changes in the A95 measure of postural sway.

Hormones

Melatonin . The group mean values for salivary melatonin levels are shown in Figure 19 in picograms per milliliter (pg/ml). Generally, melatonin was lower during watchstanding on the A Schedule than for watchstanding on the other two Schedules. Within the A Schedule, melatonin was higher during days 1 and 4, when the watches occurred at 1200-1800 and 0000-0600. Melatonin was generally low during baseline and recovery days, except for the final recovery day. To examine the reliability of these effects, the data were reduced within subjects to a mean daily peak height for the first and second 72-h Periods (three peaks each) within each of the three Schedules, then subjected to a 3-Schedule x 2-Period, repeated-measures ANOVA. There was a statistically-significant, interactive effect on mean peak

height (F(2,16) = 5.56, H-F MSe = 3.48, H-F p = 0.015) and a statistically-significant main effect of Session (F(2,16) = 8.19, H-F MSe = 17.09, H-F p = 0.005). The main effect of Period was not significant.

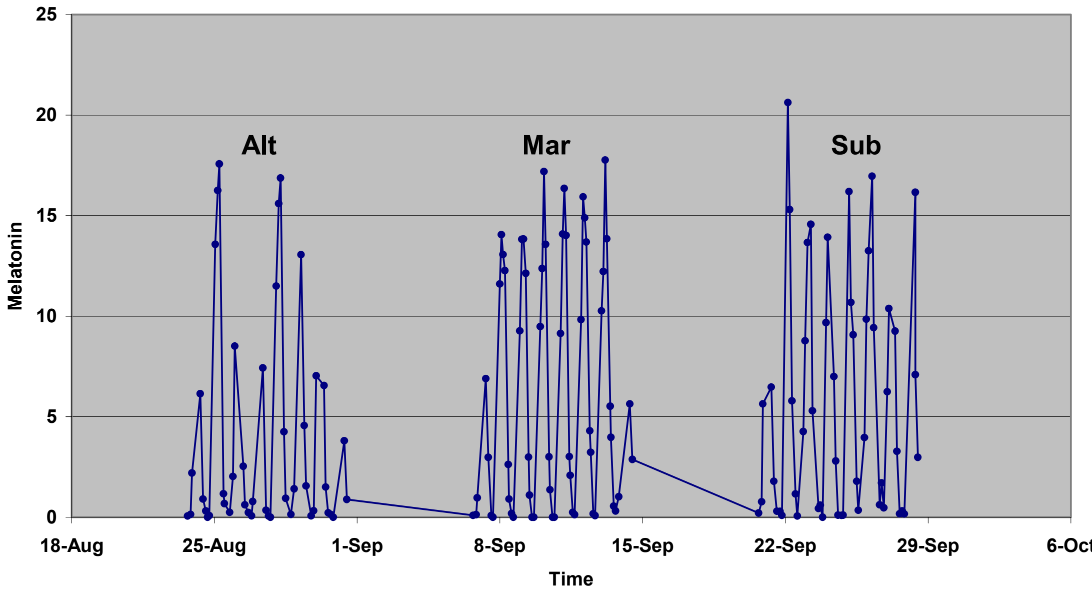

Figure 19. Group mean values for salivary melatonin in picograms per milliliter (pg/ml).

The results of the *post hoc* analysis (Figure 20) indicated that significantly lower peak levels of melatonin were present during the first period of the A Schedule than during the second Period of the A Schedule and all subsequent Periods and during the second period of the A Schedule than during all subsequent Periods. It also indicated that peak melatonin levels were significantly higher during the second period of the M Schedule than during both Periods of the A Schedule, the first Period of the M Schedule and the second Period of the S Schedule. We interpreted this pattern to indicate that lower levels of sleepiness were more likely to occur during waking periods on the A Schedule than on the other two Schedules, and that the greatest sleepiness was likely during waking periods on the M Schedule.

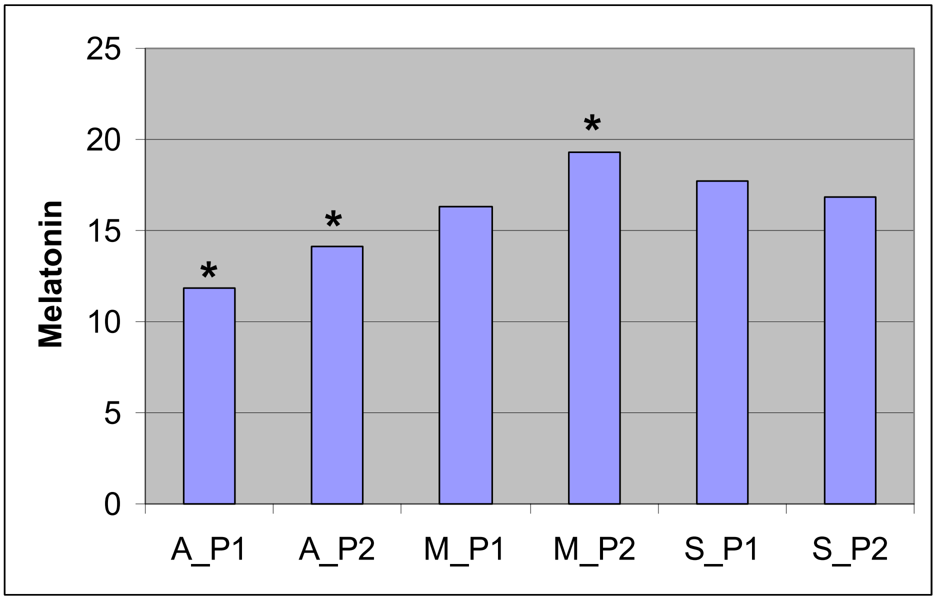

Figure 20. Mean Period peak values for salivary melatonin in picograms per milliliter (pg/ml).

Cortisol. The salivary cortisol data were reduced within subjects to a mean value for the first and second 72-h Periods within each of the three Schedules, then subjected to a 3-Schedule x 2-Period, repeated-measures ANOVA. There was a marginally significant, interactive effect on mean peak height ($F(2,13.0) = 3.56$, H-F MSe = 0.0015, H-F $p = 0.066$) and a statistically-significant main effect of Period ($F(1,8) = 7.59$, H-F MSe = 0.00088, H-F $p = 0.025$). The main effect of Session also marginally significant ($F(2,10.2) = 3.34$, H-F MSe = 0.0094, H-F $p = 0.089$).

The results of the *post hoc* analysis (Figure 21) indicated that significantly higher levels of cortisol occurred during the second Period of the A Schedule than during any other Period, and during the first Period of the A Schedule than during the first Period of the S Schedule. The highest individual reading recorded was about 2 ug/dl. Thus, the means and range of the readings were consistent with normal values for salivary cortisol reported generally in the literature. We interpreted the pattern we observed to indicate that higher levels of mild arousal were more likely to occur during waking periods on the A Schedule than on the other two Schedules.

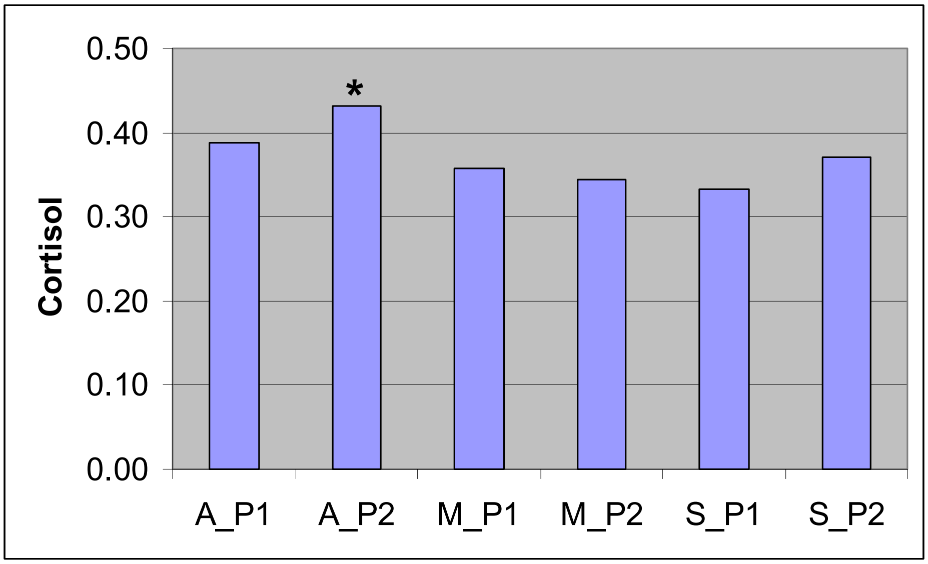

Figure 21. Mean Period values for salivary cortisol in micrograms per deciliter (ug/dl).

COSINOR ANALYSES

One measure of performance (PVT Mn_RRT), one of subjective state (SSS) and one of physiology (oral temperature) were subjected to cosinor analyses. The data of the second 72-h period (P2) were represented as a multi-day time series. Any rectilinear trend was estimated by the least squares method, described and then subtracted from the time series. The residuals were subjected to 24-h cosine curve fits using methods described by Naitoh *et al.* (1985) and Koukkari *et al.* (1974) and by Faure *et al.* (1990). The estimates for the linear regression mean and the mesor (***m****idline* ***e****stimating* ***s****tatistic* ***o****f* ***r****hythm*; the midpoint of the cosine wave, similar to a mean level) were identical. The hypothesis (h_1) that a fixed-period, 24-h cosine function fitted the residual time series was accepted if $p < 0.20$. Subsequently, acrophase and peak-to-peak (p-p) amplitude values were assessed using a single-factor, 3-level (Schedules) analysis of variance (ANOVA) with repeated measures, without randomization of orders of presentation. Tests were corrected for sphericity errors by the Huyn-Feldt procedure. Significance was accepted at $p < 0.05$. The results of the analyses are summarized in Table 8.

Psychomotor Vigilance Task Mean Reciprocal Response Time

The main effect of Schedule was statistically significant on both acrophase ($F(2,16) = 18.22$, H-F MSe = 43.557, H-F p = 0.000) and amplitude ($F(2,16) = 5.30$, H-F MSe = 0.06222, H-F p = 0.017). The *post hoc* test for acrophase indicated that it occurred significantly earlier during the M Schedule than during the A and S schedules (A, S >> M; $p < 0.01$). The *post hoc* test for amplitude indicated it was significantly smaller during the S Schedule than during the A and M schedules (A, M >> S; $p < 0.05$).

Oral Temperature

The main effect of Schedule was statistically significant on acrophase (F(2,16) = 7.81, H-F MSe = 7.328, H-F p = 0.006). Again, the *post hoc* test for acrophase indicated that it occurred significantly earlier during the M Schedule than during the A and S schedules (A, S >> M; p < 0.01). The main effect of Schedule on amplitude was not significant.

Stanford Sleepiness Scale

The main effect of Schedule was statistically significant on acrophase (F(2,16) = 11.07, H-F MSe = 12.537, H-F p = 0.001) and amplitude (F(2,16) = 7.26, H-F MSe = 1.632, H-F p = 0.015). The *post hoc* test for acrophase indicated that it occurred significantly later during the M Schedule than during the A and S schedules (M >> A, S; p < 0.01). The *post hoc* test for amplitude indicated it was significantly greater during the M Schedule than during the A and S Schedules (M >> A, S; p < 0.05).

Table 8. Summaries of cosinor analysis results for PVT performance, oral temperature and subjective sleepiness data (means +/- sd) during P2. *p < 0.05, **p < 0.01 across Schedules.

	Alternate	**Maritime**	**Submarine**
Acrophase (h)			
Performance	21:16 +/- 2.7	**06:26 +/- 7.2	19:42 +/- 4.6
Temperature	18:16 +/- 3.0	**13:43 +/- 3.0	17:41 +/- 1.4
Sleepiness	**09:16 +/- 3.9	16:05 +/- 1.8	**09:18 +/- 3.6
Amplitude (p-p)			
Performance (Mn_RRT)	0.52 +/- 0.28	0.65 +/- 0.34	*0.27 +/- 0.10
Temperature (deg F)	0.85 +/- 0.14	0.77 +/- 0.49	1.08 +/- 0.22
Sleepiness (rating)	1.28 +/- 0.41	*2.68 +/- 1.77	0.86 +/- 0.77

Reading from left to right and then top to bottom in Table 8, the circadian patterns are as follows. We predicted that the circadian acrophase in performance would remain fairly stable around 17:00 in the A Schedule (Figure 1). In fact, performance apparently phase-delayed in the A Schedule. The performance acrophase advanced in the M Schedule much farther than we had predicted (Figure 2). We predicted about a 1-hour acrophase delay in performance during the S Schedule (Figure 3), and observed about a 2-h phase-delay.

Performance usually, though certainly not always, tends to track the phase of body temperature. The temperature peak phase-delayed about an hour in the A Schedule, not nearly as far as performance. In the M Schedule, it phase-advanced but, again, not nearly as far as performance. Finally, in the S Schedule, it phase-delayed very slightly, but not as far as performance. Generally, performance acrophase shifted more than body temperature. The reason for this difference was not immediately obvious.

The sleepiness ratings presented an unusual acrophase pattern. One might expect these ratings to be approximately 180 deg out of phase with body temperature. They were not. The acrophase time for sleepiness in the M Schedule coincided nicely with the beginning of the regular, daily sleep period for that schedule (16:00 to 20:00). However, during the A Schedule, there were five sleep periods, all of which began between 12:00 and 00:00, with

the longest (10 h) starting at 22:00. During the S Schedule, there were four 6-h sleep periods: two that started in the morning (03:00 and 09:00) and two that started in the evening (15:00 and 21:00). Thus, there were no obvious relationships between sleep start time and subjective sleepiness during these two Schedules. The reasons for these 09:00 peaks in sleepiness were not immediately obvious.

Flattening of the circadian rhythm in performance and body temperature had been reported by Colquhoun *et al.* for schedules like the M Schedule. We observed flattening of the performance rhythm in the S Schedule, compared to the A and M Schedules.

Compared to the M Schedule, the amplitude of the circadian rhythms in subjective sleepiness appeared to be flattened in the A and S Schedules. It is likely that the fixed time of day for the sleep period in the M Schedule had some impact on this pattern.

CONCLUSIONS AND RECOMMENDATIONS

One major thrust of this laboratory research effort was to determine whether a sea trial of the Alternate schedule might be appropriate. Any recommendation concerning this question must take into account the larger context of shiftwork scheduling aboard ships and in other geographically-confined, limited-crew-number situations. As Schaefer *et al.* (1979) indicated, the submarine fleet apparently adopted the 12-and-6 schedule as a result of operational experiences gained during the 1960s in the early Polaris patrols. Using the traditional, 8-and-4 schedule, the crew was acquiring only 5 to 6 h/day of sleep due to additional duties and qualification training that often filled one of their two 8-h off periods. While the 8-and-4 schedule had served the needs of the maritime community for some 700 years, the piper of highly demanding technology may have played the swan song for the usefulness of this watch schedule, at least in high-technology operations.

The crews moved to the 12-and-6 schedule and immediately enjoyed the benefits of regular, 8-h periods of uninterrupted sleep. Investigations of sleep physiology conducted during the 1960s and subsequent decades indicated that this is the best way to operate for maximum human effectiveness: generally, people need about 8 h of uninterrupted time in bed to function at their best; some more, and some less. However, investigations during that same period have shown, also, that (1) the best recovery seems to occur when the time spent in bed occurs during nighttime on the body clock, and (2) the body clock can be disrupted easily by abnormal, external time cues (*Zeitgebers*), leading to fatigue, sleepiness, low motivation, feelings of malaise, etc.

Of course, the 12-and-6 watchstanding schedule (1) causes sleep to occur at various times of the 24-h cycle, and (2) provides external cues that disrupt the body clock. Undoubtedly, many crewmembers operating on the 12-and-6 schedule suffered from feelings of malaise similar to those associated with circadian rhythm disorder . Thus, the concerns of Stolgitis (1969), Johnson and Naitoh (1974), Schaefer *et al.* (1979), and Kelly *et al.* (1996) about the effects of the 12-and-6 schedule on crew performance and well-being.

It does seem that there is nothing new under the sun. The sleep length problem aboard submarines was visited in a matched set of a laboratory study and a sea trial in 1949, supported by the Naval Medical Research Institute. However, in that set of studies, the laboratory study followed the sea trial. Dr. Nathaniel Kleitman, to be hailed in subsequent decades as the father of sleep research, was a member of the Committee on Undersea Warfare of the National Research Council (Kleitman, 1949; Kleitman and Jackson, 1950; Utterback and Ludgwig, 1949). In that role, he undertook observations of the traditional 8-and-4 watch schedule aboard the USS DOGFISH.

Subsequently, Kleitman suggested that, within each 24-h period, the watches should occur closer together, thus expanding the contiguous time-off period within the 24 hours (*ibid.*). This is a zero-sum game in which 8 hours must be worked each 24 hours. It is the placement of the work and rest hours within that 24-h period that is the secret to obtaining an uninterrupted 8-h period for sleep. To reach this “close” watch solution, Kleitman broke the

eight hours of work into three periods of 3, 3 and 2 h each instead of two periods of 4 h each. His objectives were as follow. They are relevant today, just as they were in 1949.

a. Increased alertness and efficiency as a result of adjustment of working hours so that maximum body temperature might more easily coincide with them.
b. Provision for 10-12 continuous hours off, during which long uninterrupted sleep may be secured.
c. Watches of shorter duration, still providing for a total of eight hours' watch for each man.
d. A schedule so nearly impartial that the watch periods might be fixed for each section throughout the cruise.
e. A hot meal offered the men on each section before beginning their first watch of the day.
f. A dinner hour so arranged as to make it possible for the men of all sections to eat their principal and best balanced meal of the day without disrupting sleep or breaking into a watch period. (*ibid.*)

Utterback and Ludgwig (1949) and Kleitman (1949) took this concept to sea in May of 1948. They obtained oral temperature data and preferences and opinions from crewmembers during four cruises with the same crew: two cruises using the traditional Maritime 8-and-4 schedule (3.5 and 19 d), and two cruises using the proposed close watch schedule (4 and 21 d). The shorter cruises were for training, and the longer were patrols. Sample sizes varied from 15 to 28 across the four cruises. Utterback and Ludgwig concluded that "from the standpoint of sound physiology the proposed [close] schedule is superior." The crewmembers felt that the close schedule would be an improvement. Kleitman noted that schedules longer and shorter than 24 hours should be investigated, suggesting a 30-h cycle. Subsequent research has indicated that humans do not acclimatize well to non-24-h cycles.

Subsequently, Kleitman and Jackson (1950) observed nine recruits during two control periods and six periods of watchstanding of two to four weeks each, housed in the Diving Building of the Institute throughout April to August of 1949. The four watch schedules were:

1. The 8-and-4 Maritime watch schedule. The study participants represented the 0000 to 0400/1200-1600 watch section.
2. The 8-and-4 Maritime watch schedule. The study participants represented the 0800 to 1200/2000-0000 watch section.
3. The dogged watch. This is an 8-and-4 schedule in which the 1600-2000 watch is "dogged" by a different section each day, resulting in 24 h of watchstanding per 96 h instead of the 24 h in 72 h expected in the straight 8-and-4. The study participants represented only one of three possible sections.
4. The close watch schedule. The study participants represented the 0800 to 2000 watch section.
5. The close watch schedule. The study participants represented the 0000 to 1200 watch section.
6. The close watch schedule. The study participants represented the 1200 to 0000 watch section.

The investigators' objectives were (1) to detect variations in alertness as measured by simulated instrument flight performance in a Link trainer and by performance on a choice response time task and a color-naming task, and (2) to observe variations in body temperature, coffee consumption and sleep and wake times. Generally, they observed diurnal (*i.e.,* circadian) variations in body temperature and performance (note that the study of chronobiology did not evolve until the late 1950s). They also noted a tendency for better performance to be associated with a higher body temperature and with a higher volume of coffee consumption. Interestingly, they drew no major conclusions about differences in schedules. I suspect that this was because (1) they allowed the participants to sleep at least 8 h/day, and (2) the 24-h cycle was maintained on all schedules. They did note that the greatest sleep lengths (10.1 h) occurred between 0000 and 1200 and that recreational opportunities tended to interfere with time spent in bed.

Apparently, there were logistic or other hurdles, or resistance to innovation that could not be overcome to implement the close watch as a general practice. One problem may have been the shortness of the watch periods; in other words, one of the schedule's objectives may have been part of its downfall. Generally, people seem to prefer getting their work over with once they have reported for duty; they tend to prefer longer to shorter watch periods.

The present investigation introduced a "close," or "compressed," watch system using a 6-h watch period. Thus, it echoed to a large degree the suggestion made by Kleitman about 50 years before the inception of this investigation. One hopes that, if preferred by crews in sea trials, the compressed-6 schedule does not suffer the same ignominy as the close-4 schedule.

The results of the present investigation suggested that a sea trial of the compressed-6 schedule is probably warranted. To place that statement in context, understand that we have 700 years of experience with the 8-and-4 Maritime schedule and more than 30 years of experience with the 12-and-6 Submarine schedule, but just 6 d of the compressed-6 schedule. Further investigations are needed. In addition, the results of this investigation indicated that, at least for 6 d, operations on the compressed-6 schedule would "do no harm" to submariners, compared to operations on the other two schedules. Thus, if an alternative to these two older schedules is to be sought, then a sea trial of a schedule such as the compressed-6 is indicated. If the objectives of changing the watchstanding schedule from the 12-and-6 are to (1) move submariners back to a work-rest cycle of 24 h instead of 18 h, while (2) also allowing long, uninterrupted sleep periods on most days, then the options are somewhat limited, and a leading option is the compressed-6 schedule that was examined here.

Another major thrust of this investigation was to inform the Air Force R&D community about work schedule options that might be employed when 24-h operations are required in geographically-confined, limited-crew-number situations. For example, if an intelligence-gathering team must be kept as small as possible, be inserted into a confined area and operate 24 h/day, then how should their work and rest be scheduled to assure the highest possible level of job performance and lowest possible probability of job errors? This is the same work-rest scheduling problem faced by a ship's captain when underway. Thus, the conclusions drawn here that were relevant to the solution of watchstanding schedule

questions were equally relevant for USAF work-rest schedule questions of the type of described here.

SPECIFIC CONCLUSIONS

There was one obvious, detectable effect of Schedule on performance: it appeared that the M Schedule provided about a 10% advantage in response time on the PVT over the A Schedule. However, because of the paucity of similar analytic results for other tasks, and for other measures acquired from this task, we were hesitant to draw conclusions about experimental Schedule effects based solely upon this one difference. This unsupported occurrence of statistical significance may well have been due to random, not experimental, effects.

Often, there is as much knowledge to be gained by noting things that do not happen as by studying the patterns of things that do happen. That was certainly true with regard to the performance data acquired in this investigation. We do know that the some of the performance measures we used were sensitive to the effects of learning, recovery (Running Memory task omissions) and Periods. Thus, the lack of Schedule effects on performance was quite interesting. Apparently, through the first six days, one schedule is as good as the other in terms of performance. Colquhoun *et al.* noted that the time required for performance changes to become obvious may be more than four 72-h cycles, though they had detected some differences in the first two cycles.

The subjective measures we acquired presented a mixture of findings. During P2, Mood 2-R Activity ratings were highest before watches in the A and S Schedules, and lowest after watches in the A Schedule. The latter effect on the A Schedule was likely due to the compression of 12 h of work into 18 h in the A Schedule. The lesson here is that there is no free lunch: compressing work to expand contiguous time off carries with it the burden of a greater work rate. For example, 12 hours of work in 16 h meant working 75% of the time in the A Schedule. However, 4 h of work in 12 h or 6 h of work in 18 h meant working 33% of the time in the M and S schedules, respectively. At the end of the compressed 12 h of work in the A Schedule, the participants did not feel very energetic, lively, alert, spirited, active, or steady.

Mood 2-R Fatigue ratings were highest in the M Schedule and lowest in the A Schedule. This effect may have represented the predicted presence of the malaise associated with circadian rhythm disorder in the M Schedule, contrasted with the expected circadian stability of the A Schedule.

During P2, there were increased reports of the symptoms, "Trouble staying awake" and "'Drugged' feeling" during the A Schedule. These findings are in conflict with the Mood 2-R Fatigue effects, above. They may have been associated more with the facts that the A Schedule was worked first and that it was quite unfamiliar to the participants than with perceptions of impairment due solely to circadian rhythm and/or sleep disruptions. Alternatively, these responses may have reflected the fragmentation of sleep periods in the A Schedule.

The somatic pre-sleep score for Recovery after the A Schedule was significantly greater than for Recovery after the S Schedule   This effect could have been associated with greater anxiety following participation in the A Schedule than the familiar S Schedule. However, there were no related (such as pupil dilation) data to support that contention. More likely, the effect was due to the lesser need for recovery sleep after the A Schedule.

Thus, though the signals were mixed, the main picture that arose from the subjective measures suggested that:

1. The compressed work pattern of the A Schedule had a noticeable effect on perceptions of Mood 2-R "Activity" that should be examined in a longer study. The effect noted here may have been a precursor of the development of cumulative fatigue.
2. The malaise predicted to occur as a result of circadian rhythm disorder caused by the M Schedule apparently surfaced as a perception detected by the Mood 2-R "Fatigue" scale.

We learned more from the physiological measures employed in this study, especially polysomnography (PSG), than from the performance or subjective measures. The participants achieved the expected average 7.0 h TST on all three Stabilization nights. Their sleep was not quite as efficient as expected, but adequate. Subsequently, P1 was characterized by many variations in sleep patterns. Some of these were probably induced by the circadian rhythm disorder associated with the M Schedule: sleep efficiency was poor due to high WASO in the M Schedule. Others may have been induced by the compressed sleep periods (6 h TIB) of the S Schedule: %S2 was relatively low and %S4 and %SREM relatively high compared to the other schedules. The M Schedule, during which sleep was compressed to 7 h TIB, showed these same tendencies.

By P2, sleep patterns had settled down. The A Schedule was the only Schedule in which the participants achieved the expected 7.0 h of TST. They acquired averages of 2.7 and 4.0 hours more sleep in P2 during the Alternate schedule than during P2 of the Maritime and Submarine schedules, respectively.

The sleep patterns displayed during Recovery suggested that the Alternate schedule created a lesser need for slow-wave sleep (S3) than the other two schedules, and that stage 2 sleep may have substituted for slow-wave sleep after the Alternate watchstanding schedule. Of course, the participants had just spent 24 h off from watchstanding before the Recovery period after the Alternate schedule. During that period, their final sleep was 10 h long (2200 to 0800) and then they were awake for 10 h (0800 to 1800) before recovery sleep. In comparison, at the end of the Maritime schedule the participants' final sleep was slept only 7 h long (1630 to 2330) and they were awake for 18.5 h (2330 to 1800) before recovery sleep. At the end of the Submarine schedule the participants' final sleep was only 6 h long (0800 to 1800) but they had also been awake for 10 h (2330 to 1800) before recovery sleep.

Subsequently, after the R Period, the participants generated more sleep (as estimated by actigraphy) after the M Schedule than after the A or S Schedules, suggesting that a greater degree of recovery was required after the M Schedule than after the A Schedule. This conclusion was supported further by the mean peak values for salivary melatonin. These suggested that lower levels of sleepiness (lower mean peak levels of melatonin) were more

likely during waking periods on the A Schedule than on the other two Schedules, and the occurrence of the greatest level of sleepiness (highest mean peak levels of melatonin) was most likely during waking periods on the M Schedule. The conclusion was supported further by the mean values observed for salivary cortisol. These suggested that higher levels of mild arousal were more likely to occur during waking periods on the A Schedule than on the other two Schedules.

The main picture that arose from the PSG and hormone measures suggested that:

3. As the participants adjusted to the new work-rest schedules during P1, sleep efficiency and structure was mildly affected by circadian rhythm disorder and short sleep periods.
4. During P2, more good-quality sleep was acquired in the A Schedule than in the other two schedules.
5. The need for Recovery sleep was not an issue following the A Schedule, perhaps underscoring a greater compatibility of the A schedule with normal human biology. It was definitely an issue following the M Schedule—enough cumulative fatigue had built up during the M Schedule that recovery took more than 24 h.

Specific examinations of circadian rhythms in body temperature, simple unalerted response time performance and perceived sleepiness supported the advantages of the fixed work-rest schedule. The performance circadian rhythm acrophase advanced in the M Schedule much farther than we had predicted. It also delayed to some degree in the other two Schedules. In all three Schedules, performance acrophase tended to shift more than body temperature. The reason for this difference was not immediately obvious.

The acrophase time for sleepiness in the M Schedule coincided nicely with the beginning of the regular, daily sleep period for that schedule (16:00 to 20:00). The reasons for the 09:00 peaks in sleepiness in the other two Schedules were not immediately obvious.

Compared to the M Schedule, the amplitude of the circadian rhythms in subjective sleepiness appeared to be flattened in the A and S Schedules. It is likely that the fixed time of day for the sleep period in the M Schedule had some impact on this pattern. There was also a flattening of the performance rhythm in the S Schedule, compared to the A and M Schedules.

Thus, the main picture that arose from the circadian rhythm analyses suggested that:

6. The participants adjusted their body clocks quickly to the fixed work-rest schedule of the M Schedule with a quick advance of their performance and body temperature acrophases, alignment of the sleepiness peak with the regular daily bedtime and a Sustainment of the strength (amplitude) of the normal circadian rhythm in metabolism (body temperature).

Finally,

7. The high pre-sleep somatic scale score values and the significantly larger pupil size associated with P2 of the M Schedule were consistent with generalized, elevated sympathetic tone and/or generalized, decreased parasympathetic tone. These findings reflected two obvious effects on the participants of the terrorist acts of September 11, 2002.

SPECIFIC HYPOTHESES

24-hour work-rest cycles will produce better entrainment of circadian rhythms in physiology and performance to the 24-hour clock than will an 18-hour work-rest cycle. This hypothesis was supported by Conclusion 6, that the participants adjusted their body clocks quickly to the fixed work-rest schedule of the M Schedule.

Given the same average amount of time in bed and average time spent on watch per 24 hours, both sleep quality and sleep quantity will be worse in an 18-hour work-rest schedule than in 24-hour work-rest cycles. This hypothesis was supported by Conclusion 4, that more good-quality sleep was acquired during P2 in the A Schedule than in the other two schedules.

Given the same average amount of time in bed and average time spent on watch per 24 hours, performance and mood will be worse in an 18-hour work-rest schedule than in 24-hour work-rest cycles. This hypothesis was not supported.

Given the same average amount of time in bed and average time spent on watch per 24 hours, both sleep quality and sleep quantity will be worse in a standard maritime work-rest schedule than in an alternative 24-hour work-rest cycle. This hypothesis was supported by Conclusions 4 and 5, that that more good-quality sleep was acquired during P2 in the A Schedule than in the other two schedules, and that the need for Recovery sleep was not an issue following the A Schedule but was definitely an issue following the M Schedule.

Given the same average amount of time in bed and average time spent on watch per 24 hours, performance and mood will be worse in a standard maritime work-rest schedule than in an alternative 24-hour work-rest cycle. The mood portion of this hypothesis was supported by Conclusion 2, that the malaise predicted to occur as a result of circadian rhythm disorder caused by the M Schedule apparently surfaced as a perception detected by the Mood 2-R "Fatigue" scale. The performance portion of the hypothesis was not supported.

LESSONS LEARNED

Since the late 1970s, fatigue investigators have tended toward a 3-pronged approach to fatigue measurement: we measure performance and physiology and we quantify self-reports of perceived fatigue. Usually, we learn different, but related, things from these three kinds of measures. The data acquired from this study are a case in point. The performance measures were not sensitive to the schedule manipulations in this experiment, but we learned several different, but related things from the other two measurement domains. Why were the performance measures insensitive?

Performance is often the "bottom line" when we are concerned about crew fatigue[8]. We are often asked to predict whether a crew may accomplish a given duty period safely. This is a reasonable question. Most often, the answer is "Most of the time." After all, aircrews make many, many flights in very fatigued states and do not have accidents.

---

[8] This and the next paragraph were adapted from Miller JC (2001), *Controlling Pilot Error: Fatigue*, McGraw-Hill.

However, one's overt performance is not always sensitive to the effects of fatigue. This problem is due to the "two-edged sword" of human adaptability. The "good" edge is the ability of military personnel to motivate themselves to face challenges and to accomplish difficult tasks in acceptable manners in the presence of high levels of strain and resulting fatigue. Typically, the fatigued but motivated human can mobilize his or her physical and cognitive resources quite well for brief periods. This is the "can-do" attitude, characterizing what we think of as good crews.

The "bad" edge of the sword is the eventual effect of physiological and mental costs: there may be a cessation in performance (a mental lapse) or an involuntary onset of sleep (falling asleep on the job). The measured performance of the fatigued but motivated crewmember may show no impairment at all until performance ceases abruptly. Thus, fatigue is a covert result of the costs generated by effort and performance.

In the present investigation, we observed covert signs of fatigue in sleep measures and subjective reports. These covert signs may have been precursors of more overt signs of fatigue, including performance impairments that would have occurred after several more 72-h cycles, as observed by Colquhoun *et al.* The natural extension of this investigation, then, is a lengthier study or a lengthier sea trial.

OTHER ALERNATIVES

The compressed-6 (A) schedule applied the "close" watch, or compression, technique suggested by Kleitman in 1949. It differed from the original close watch schedule by extending the watch period from 4 h to 6 h, as used in the existing 12-and-6 S Schedule, instead of shortening the watch to 2 and 3 h. The arithmetic benefit of this difference was the provision of a period of 24 h of rest every third day. However, both compressed schedules induced fatigue due to compressed work and fragmented, irregular sleep periods. The compressed schedules may be improved by changing them to a fixed work-rest cycle with less sleep fragmentation.

One should consider the most extreme 8-h-work schedule compression that may be achieved within a 24-h period: an 8-and-16, fixed, close watch schedule. Work by the leading scientists, Kleitman, Colquhoun, Aschoff, and Wever, and their colleagues in various studies supported the need for a regular, long, protected sleep period for the performance and well-being of watchstanders. (Of course, Colquhoun's work also supported the usefulness of a rotating watch schedule rather than a fixed watch schedule if one wishes to avoid the painful period of adjustment to a new, fixed schedule as the boat gets underway. However, in the long run, this objective is less important for watchstanders on a long cruise than the benefits to be accrued from the fixed schedule.) Colquhoun, in fact, recommended using a 1-in-3, 8-h fixed watch schedule (Colquhoun, *et al.*, 1978; Colquhoun, 1985). Alternatively, he recommended a modification of Kleitman's close watch schedule that used fixed work and rest periods. Relevant to a consideration of using a long, fixed work period, recent findings from another study suggested that sailors on a 12-on, 12-off work-rest schedule acclimatized more fully to an aircraft carrier's night-work, day-sleep schedule than sailors who stood 4- to

6-h watches[9].

Unfortunately, the 12-and-12 becomes a rotating schedule unless either two or four teams are used. Changing the Navy from a 3-team approach to two or four teams is unlikely. However, these options are valid for 24-h operations in geographically-confined, limited-crew-number situations. Even more unfortunately, this same arithmetic applies to the A Schedule when one attempts to change it to a fixed-period schedule.

Another approach to the watchstanding problem that combines the idea of watch compression with a fixed work-rest schedule is a dogged 6-h watch in which one of the four 6-h quadrants of the day is dogged by three teams in 2-h segments[10]. In the following example (Figure 19), the 12:00-18:00 watch period is shown as the dogged watch but, in theory, any one of the other three watch periods could be dogged, as well.

Team	00-06	06-12	12-14	14-16	16-18	18-00
1	X		*x*			
2		X		*x*		
3					*x*	X

Figure 22. An example of a dogged-6 watch schedule.

This kind of approach provides for 8 h of watch per day, as is the current practice. This particular schedule allows the following time-off periods:

- Team 1, 06:00-12:00 (6 h) and 14:00-00:00 (10 h); afternoon and evening sleep
- Team 2, 12:00-14:00 (2 h) and 16:00-06:00 (12 h); overnight sleep
- Team 3, 00:00-16:00 (16 h); overnight-early morning sleep; 8 h continuous watch (16:00-00:00)

Thus, several alternatives to the A, M and S Schedules present themselves for sea trials:

1. The fixed, close watch schedule (Colquhoun)
2. A fixed, dogged-6 schedule
3. The fixed, 8-h watch schedule (Colquhoun)

Several alternatives to the A, M and S Schedules present themselves for 24-h operations in geographically-confined, limited-crew-number situations where 2 or 3 teams may be used:

1. The fixed, close watch schedule (Colquhoun)
2. A fixed, dogged-6 schedule
3. The fixed, 8-h watch schedule (Colquhoun)
4. A fixed version of the A Schedule
5. The fixed, 12-and-12 schedule

RECOMMENDATIONS

[9] Dr. N. Miller, Naval Postgraduate School, personal communication, February 2002.

[10] The initial "dogging" concept by the British Navy was aimed at causing rotation of the watchstanding schedule. Subsequently, the term became a descriptor of the simple splitting of one watch period across watch bills, whether rotation was a result or not.

If, in fact, the work compression and expansion of time off that may be achieved by lengthening the watch is desirable, then the following 3-team schedules should be considered for sea trials:

- The A Schedule (compressed-6)
- A fixed, dogged-6 schedule
- The fixed, 8-h watch schedule

For 24-h operations in geographically-confined, limited-crew-number situations consider:

- The A Schedule (compressed-6)
- A fixed version of the A Schedule
- A fixed, dogged-6 schedule
- The fixed, 8-h watch schedule
- The fixed, 12-and-12 schedule

If work compression and expansion of time off is to be achieved by shortening the watch, then the following schedule should be considered for sea trials and for 24-h operations in geographically-confined, limited-crew-number situations:

- The fixed, close watch schedule

One or more of the intake tools used in this study may be useful in the selection process for shift and night workers. As noted, above, morning types may report low satisfaction with night work and may opt out of shift and night work. Conversely, evening types may tend more toward acceptance of night and shift work and thus may be the people who often go on to develop the kinds of health problems generally associated with night and shift work (Kundi *et al*., 1986). Thus, the morningness-eveningness questionnaire used here may provide added value in the selection process by predicting attrition and/or health problems.

REFERENCES

Alluisi EA (1967). Methodology in the use of synthetic tasks to assess complex performance. *Human Factors*, 9, 375-384.

Ames LL, George EJ (1993). *Revision and Verification of a Seven-Point Workload Estimate Scale* (AFFTC-TIM-93-01). Edwards AFB CA, Air Force Flight Test Center.

Arendt J, Stone B, Skene D (2000). Jet lag and sleep disruption. In MH Kryger, T Roth, WC Dement (ed.), *Principles and Practice of Sleep Medicine*, W.B. Saunders, Philadelphia.

Aschoff J, Fatranska M, Giedke H *et al.* (1971). Human circadian rhythms in continuous darkness: entrainment by social cues. *Science*, 171, 213-215.

Beck AT, Ward, C H, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. *Archives of General Psychiatry* 4, 561-571.

Borg, G. (1985). *An Introduction to Borg's RPE-Scale*. Ithaca NY, Movement Publications.

Caille EJ, Bassano JL (1977). Biorhythm and watch rhythms: Hemeral watch rhythm and anhemeral watch rhythm in simulated permanent duty. Chapter 22 in RR Mackie (ed.), *Vigilance: Theory, Operational Performance and Physiological Correlates*. Plenum Press, New York.

Chapman S (2001). *Fatigue Management amongst Royal Australian Navy Submariners.* Report submitted for the Submarine Service Allowance Review.

Chiles WD, Alluisi EA, Adams OS (1968). Work schedules and performance during confinement. *Human Factors* 10(2), 143-196.

Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale NJ.

Cole RJ, Kripke DF, Gruen W, Mullaney DJ, and Gillin CJ (1992). Automatic sleep/wake identification from wrist activity. *Sleep*, 15(5), 461-469.

Colquhoun WP (1980). Simulated shiftwork. In WP Colquhoun, J Rutenfranz (ed.), *Studies of Shiftwork*, Taylor & Francis, London.

Colquhoun WP (1985) Hours of shiftwork at sea: watchkeeping schedules, circadian rhythms and efficiency. *Ergonomics*, 28(4), 637-653.

Colquhoun WP, Blake MJF, Edwards RS (1968). Experimental studies of shift-work I: A comparison of 'rotating' and 'stabilized' 4-hours shift systems. *Ergonomics* 11, 437-453.

Colquhoun WP, Paine MWPH, Fort A (1978). Circadian rhythm of body temperature during prolonged undersea voyages. *Aviat. Sp. Environ. Med.* 49(5), 671-678.

Colquhoun WP, Paine MWPH, Fort A (1979). Changes in the temperature rhythm of submariners following a rapidly rotating watchkeeping system for a prolonged period. *Int. Arch. Occup. Environ. Health* 42, 185-190.

Dinges DF (1992). Probing the limits of functional capability: The effects of sleep loss on short-duration tasks. In Broughton, R.J.and Ogilvie, R.D. (Eds.), *Sleep, Arousal, and Performance*, Boston, Birkhäuser, pp. 176-188.

Dinges DF, Pack F, Williams K, Gillen KA, Powell JW, Ott GE, Aptowicz C, and Pack AI (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep*, 20(4), 267-277.

Donderi DC, Smiley A, Kawaja KM (1995). *Shift Schedule Comparison for the Canadian Coast Guard.* Transport Canada Publication no. TP 12438E, Transportation Development Centre, Montreal, Quebec.

Eddy DR, Gibbons J, Cardenas R, Hickey P, Barton E, Stevens K, Michel A-M, French J, Storm WF, Miller JC (2002). *An Assessment of Modafinil for Vestibular and Aviation-Related Effects.* Technical report in preparation, Air Force Research Laboratory, Brooks AFB TX.

Elsmore TF (1994). SYNWORK1: A PC-based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instruments, & Computers*, 26, 421-426.

Englund CE, Reeves DL, Shingledecker CA, Thorne DR, Wilson KP, Hegge FW (1987). *The Unified Tri-Service Cognitive Performance Assessment Battery (UTC-PAB): I. Design and Specification of the Battery* (Report No. NHRC-TR-87-10). Naval Health Research Center, San Diego, CA..

Faure A, Nemoz C, Claustrat B (1990). A graphical and statistical method for investigation of time series in chronobiology according to the cosinor procedure. *Comput Biol Med*, 20(5), 319-329.

Fess EE, Moran CA. Cited by: Mathiowetz V, Weber K, Volland G, Kashman N (1984). Reliability and validity of grip and pinch strength measurements. J Hand Surg., 9A(2), 222-226.

Grandjean E (1982). *Fitting the Task to the Man*. London, Taylor & Francis.

Hildebrandt G (1976). Outline of chronohygiene. *Chronobiologia*, 3(2), 113-126.

Hockey GRJ, Colquhoun WP (1972). Diurnal variation in human performance: a review. In WP Colquhoun (ed.), *Aspects of Human Efficiency: Diurnal Rhythm and Loss of Sleep*, The English Universities Press, London.

Hoddes E, Zarcone VP, Smythe H, Phillips R, Dement WC (1973). Quantification of sleepiness: A new approach. *Psychophysiology,* 10, 431-436.

Horne JA (1991). Dimensions to sleepiness. Chapter 7 in TH Monk, *Sleep, Sleepiness and Performance*, Wiley, pp. 169-196.

Horne JA, Ostberg O (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int J Chronobiol*, 4(2), 97-110.

Hubinger L (1998). Great Barrier Reef Pilots Questionnaire. Queensland University of Technology, School of Human Movement Studies, Kelvin Grove, Queensland, Australia.

Johns MW (1991). A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. *Sleep*, 14, 540-545.

Johns MW (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep*, 15, 376-381.

Johnson LC, Naitoh P (1974), *The Operational Consequences of Sleep Deprivation and Sleep Deficit*, AGARD-AG-193, Advisory Group for Aerospace Research and development, North Atlantic Treaty Organization, Neuilly sur Seine, France.

Kelly T, Neri DF, Grill JT, Ryman D, Hunt PD, Dijk DJ, Mitchell J, Shanahan TL, Czeisler CA (1996). *Circadian Rhythms in Submariners Scheduled to an 18-Hour Day*. Publication 96-42, Naval Health Research Center, San Diego.

Kilbom A (1991). Measurement and assessment of dynamic work. Ch. 22 in J.R. Wilson and E.N. Corlett (Eds.), *Evaluation of Human Work*, London, Taylor & Francis.

Kleitman N (1949). The sleep-wakefulness cycle of submarine personnel. Ch. 15 in *Human Factors in Submarine Warfare*, Waverly Press, Baltimore, pp. 329-341.

Kleitman N, Jackson DP (1950). *Variations in Body Temperature and in Performance Under Different Watch Schedules.* Project 004.005.01.02, Naval Medical Research Institute, Bethesda MD. (AD-667-515)

Kobasa SC, Maddi SR (1977). Existential personality theory. In R. J. Corsini (Ed.), *Current Personality Theories*, (pp. 243-276). Itasca, IL: F. E. Peacock Publishers.

Kobasa SC, Maddi SR, Kahn S (1982). Hardiness and health: A prospective study. *Journal of Personality and Social Psychology*, 42, 168-177.

Koukkari WL, Duke SH, Halberg F, Lee J-K (1974). Circadian rhythmic leaflet movements: student exercise in chronobiology. *Chronobiologia*, 1(3), 281-302.

Kraft GH, Detels PE. Position of function of the wrist. *Arch Phys Med Rehabil*. 1972;53:272-275.

Kubo T, Sakata Y, Matsunaga T, Koshimune A, Sakai S, Ameno K, Ijiri I (1989). Analysis of body sway pattern after alcohol ingestion in human subjects. *Acta Otolaryngol* (Stockh) 468, 247-252.

Kundi M, Koller M, Cervinka R, Haider M (1986). Health and psychosocial aspects of shiftwork results of a 5-year follow-up study. In F Nachreiner (ed.), *Studies in Industrial and Organizational Psychology (Vol. 3)*. Verlag Peter Lang, Frankfurt & New York.

Lacks P (1987). *Behavioral Treatment for Persistent Insomnia*. New York: Pergamon.

Lacks P, Robert M (1986). Knowledge and practice of sleep hygiene techniques in insomniacs and good sleepers. *Behaviour Research and Therapy*, 24, 365-368.

MacDermid JC, Kramer JF, Woodbury MG, McFarlane RM, Roth JH (1994). Interrater reliability of pinch and grip strength measurements in patients with cumulative trauma disorders. *J Hand Surg*., 7, 10-14.

Mathiowetz V, Weber K, Volland G, Kashman N (1984). Reliability and validity of grip and pinch strength measurements. *J Hand Surg*., 9A(2), 222-226.

Mawdsley RH, Girio JA, Desmet EC, Bull DJ, Petrick ST, Hartman RS (1999). Intertester reliability of grip strength measurements of elderly females. *Issues on Aging*, v. 22.

Miller JC (1992). *Fundamentals of Shiftwork Scheduling*. Miller Ergonomics, Lakeside CA.

Miller JC, Smith ML, McCauley ME (1999). *Crew Fatigue and Performance on U.S. Coast Guard Cutters*. Report no. CG-D-10-99, U.S. Coast Guard Research and Development Center, Groton CT. (AD A366708)

Morris TL, Miller JC (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology*, 42, 343-360.

Naitoh P, Englund CE, Ryman DH (1985). Circadian rhythms determined by cosine curve fitting: analysis of continuous work and sleep loss data. *Behavior Research Methods, Instruments & Computers*, 17(6), 630-641.

Naitoh P, Beare AN, Biersner RJ, Englund CE (1983). Altered circadian periodicities in oral temperature and mood in men on an 18-hour work/rest cycle during a nuclear submarine patrol. *Int J Chronobiology*, 8(2), 149-173.

Nicassio PM, Mendlowitz DR, Fussell JJ, Petras L (1985). The development of the Pre-Sleep Arousal Scale. *Behaviour Research and Therapy*, 23, 266.

Patat A, Foulhoux P (1985). Effect on postural sway of various benzodiazepine tranquillizers. *British Journal of Clinical Pharmacology*, 20, 9-16.

Pressman, M.R., DiPhillipo, M.A., and Fry, J.M. (1986). Senile miosis: the possible contribution of disordered sleep and daytime sleepiness. *J. Gerontology*, 41, 629-634.

Pryce JC (1980). The wrist position between neutral and ulnar deviation that facilitates the maximum power grip strength. *J Biomech*., 13, 505-511.

Ranzijn R, Lack L (1997). The pupillary light reflex cannot be used to measure sleepiness. *Psychophysiology*, 34, 17-22.

Rechtschaffen A, Kales A (ed.) (1968). *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Public Health Service, US Government Printing Office, Washington, DC.

Reeves D, Winter K, Kane R, Elsmore T, Bleiberg J (2001). *ANAM 2001 User's Manual*. Special Report NCRF-SR-2001-1, National Cognitive Recovery Foundation, San Diego CA.

Rowland L, Krichmar J, Sing H, Thomas M, Thorne D (1997) "Pupil dynamics and eye movements as indicators of fatigue and sleepiness." *Proc. Annual Meeting of the Association for Research in Vision and Ophthalmology*, Ft Lauderdale FL.

Russo M, Thomas M, Sing H, Thorne D, Balkin T, Wesensten N, Redmond D, Welsh A, Rowland L, Johnson D, Cephus R, Hall S, Krichmar J, Belenky G (1999) "Saccadic Velocity and Pupil Constriction Latency are Sensitive to Partial Sleep Deprivation, and Deprivation Related Changes Correlate with Simulated Motor Vehicle Crashes," *presentation to American Academy of Neurology Annual Meeting*, Walter Reed Army Institute of Research, Washington DC

Rutenfranz J, Plett R, Knauth P, Condon R, DeVol D, Fletcher N, Eickhoff S, Schmidt KH, Donis R, and Colquhoun WP (1995). Work at sea: A study of sleep, and of circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels - II. Sleep duration, and subjective ratings of sleep quality. *Occupational and Environmental Health*, 60, 331-339.

Ryman D, Biersner R, Rocco J (1974). Reliabilities and validities of the mood questionnaire. *Psychol Rep*, 35, 479-484.

Sandquist TF, Raby M, Maloney AL, and Carvalhais AB (1996). *Fatigue and Alertness in Merchant Marine Personnel: A Field Study of Work and Sleep Patterns* (US Coast Guard Report no. CG-D-06-97). Washington DC, US Coast Guard. (AD-A322126)

Schaeffer KE, Kerr CM, Buss D, Haus E (1979). Effect of 18-h watch schedule on circadian cycles of physiological functions during submarine patrols. *Undersea Biomedical Research, Submarine Supplement*, 6, S81-S90.

Schmidt HS, Jackson EI, Knopp W (1981). Electronic pupillography (EPG): objective assessment of sleepiness and differentiation of disorders of excessive somnolence. *Sleep Res.*, 10, 48.

Spielberger CD (1983). *Manual for the State-Trait Anxiety Inventory (STAI).* Consulting Psychologists Press, Palo Alto, CA.

Spielman AJ, Yang C-M, Glovinsky PB (2000). Assessment techniques for insomnia. Chapter 103 in MH Kryger, T Roth, WC Dement (ed.), *Principles and Practices of Sleep Medicine (3rd ed.)*, WB Saunders, Philadelphia.

Stampi C, Aguirre A, Macchi M, Hashimoto S (1994) *Evaluation of Pulse FIT Parameters for Detection of Fatigue (Reduced Alertness),* Institute for Circadian Physiology, Cambridge MA.

Stanley BG, Tribuzi SM (1992). *Concepts in Hand Rehabilitation*. F. A. Davis Company, Philadelphia, PA.
Stolgitis WC (1969). *The Effects of Sleep Loss and Demanding Work/Rest Cycles: An Analysis of the Traditional Navy Watch System and Proposed Alternative.* Thesis S738, Naval Postgraduate School, Monterey, CA.
Umeda T, Hiramatsu R, Iwaoka T, Shimada T, Miura F and Sato T (1981). Use of Saliva for Monitoring Unbound Free Cortisol Levels in Serum. *Clin Chem Acta*, 110, 245-253.
Utterback RA, Ludwig GD (1949). *A Comparative Study of Schedules for Standing Watches aboard Submarines Based on Body Temperature Cycles.* Report no. 1, Project 004 003, Naval Medical Research Institute, Bethesda MD. (AD-667-707)
Van Cauter E (2000). Endocrine Physiology. Chapter 20 in Kryger MH, Roth T, Dement W (2000), *Principles and Practice of Sleep Medicine*, WB Saunders.
Wever RA (1979). *The Circadian System of Man*. Springer-Verlag, New York.
Wever RA (1983). Organization of the human circadian system: internal interactions. In TA Wehr and FK Goodwin (ed.), *Circadian Rhythms in Psychiatry*, Boxwood, Pacific Grove CA.
Wever RA (1989). Light effects on human circadian rhythms: a review of recent Andechs experiments. *J. Biol. Rhythms*, 4, 161-185.
Wever RA, Polasek J, Wildgruber CM (1983). Bright light affects human circadian rhythms. *Pflügers Archive*, 396, 128-137.
Wiker SF, Kennedy RS, McCauley ME, and Pepper RL (1979). *Reliability, Validity, and Application of an Improved Scale for Assessment of Motion Sickness Severity*. Technical Report No. CG-D-29-79, U.S. Coast Guard Office of Research and Development, Washington, DC.
Williams RL, Karacan I, Hursch CJ (1974). *EEG of Human Sleep: Clinical Applications*. John Wiley, New York.
Wylie CD, Shultz T, Miller JC, Mitler MM, Mackie RR (1996). *Commercial Motor Vehicle Driver Fatigue and Alertness Study: Project Report*. Technical Report FHWA-MC-97-002, Federal Highway Administration, Department of Transportation, Washington DC. (also, simultaneous Transport Canada Publication no. TP 12875E)

Appendix A

EXPANDED DISCUSSION OF PROPOSED WATCH SCHEDULES

The "Submarine" Watch Schedule

This was a 3-crew, 6-h, 2:1 rotation that placed the crewmember on an 18-hour work-rest cycle. The specific characteristics of the Submarine schedule were:

- Cycle length: 72 hours
- Watch length: 6 hours
- Number of watches: 4 in 72 hours (same number as traditional maritime watch schedule)
- Longest time off: 12 hours
- Work demand: 24 hours of watch per 72 hours (same as traditional maritime watch schedule)

The 72-hour watch sequence:

Sequence of Watches (6 h)	**Scheduled Sleep Time (6 h)**	**Predicted Best Sleep Times**
0000-0600	0900-1500	1330-1700 (3.5 h)
1800-0000	0300-0900	0030-0900 (8.5 h)
1200-1800	2100-0300	2030-0500 (8.5 h)
0600-1200	1500-2100	1330-1700 (3.5 h)

The watch time, scheduled sleep time and predicted best sleep times all sum to 24 h per 72 h, or an average of 8 h per 24 h.

The "Maritime" Watch Schedule
(Review material adapted from Miller *et al.*, 1999)

This was a 3-crew, 4-h, 2:1 rotation. The genesis of the traditional 4-h watch in marine operations, reported as early as the 13th century, is obscure. However, it is likely that it came into use because it meets several criteria: (1) 4 h is a factor of the 24-h day, (2) 4 h is a factor of an 8-h work day, and (3) 4 h is viewed as an acceptable length of time during which one can stand upright without excessive fatigue and also withstand stressful environmental factors such as wind, rain and cold.

Concerning criterion 1, being a factor of 24 h, it is quite difficult to design and maintain a work-rest schedule based upon watch lengths that are not factors of the 24-h day, such as 4, 6, 8 and 12 h (Miller, 1992). Thus, a 4-h watch length helps create a watch schedule that is simple to understand, plan and execute. As to being a factor of an 8-h workday, errors tend to increase disproportionately if one continues to perform physical labor beyond 8 h per day (cf. Grandjean, 1982). This latter observation has existed in the work research literature for decades and may have been self-evident long before that. Thus, carrying three teams on board may meet a minimum criterion for fatigue and error production: with two teams on 12 h of work per day, too many errors might occur due to fatigue. The use of four teams to

reduce fatigue effects would present other problems: with four teams working 6 h per day, or working 8 h per day and taking one day in four off, crew members would probably become bored and inefficient. Also, the boat would need to carry supplies for 1.3 times as many people as it does for three teams.

The use of three teams instead of four carries with it a relative workload penalty. In industry, an 8-h system is staffed with four crews so that one crew is in recovery (days off) at all times. Thus, a 3-crew, 8-h maritime watch system calls for more work per unit time from a watchstander, by a factor of 1.3, than the standard, industrial 4-crew, 8-h system. This relationship is non-linear (see figure, below). The penalty increases disproportionately with fewer and fewer crews performing a given amount of work. Alternatives to the traditional maritime watch schedule should not induce a further penalty.

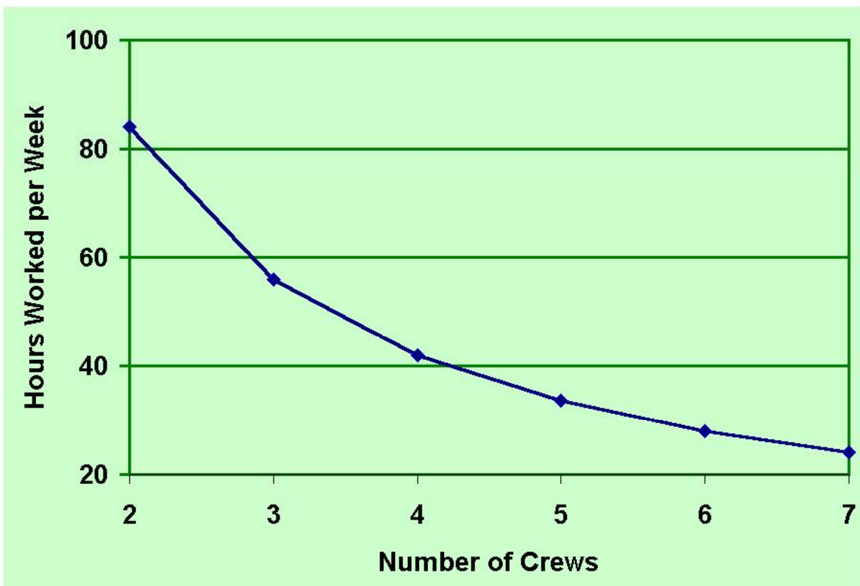

Figure A1. Average hours worked per week as a function of the number of crews available

Finally, standing a watch often entails *standing* during the entire watch. This was true for USCG cutter Quartermasters, in particular (Miller *et al.*, 1999; interestingly, helmsmen were seated in two of the six cutters observed). Certainly, there is historical precedent for seamen literally *standing* a watch. Whatever the reason, standing for four hours and then taking an 8-hour break from continual standing is, obviously, more palatable than standing for eight hours without a break. Whether or not four hours is a reasonable amount of time to stand depends in part upon the leg muscle tone and cardiovascular capabilities of an individual, but it is a physically demanding effort, particularly in high sea states.

In its favor, the 4-h watch can lend itself to a physiologically-regular work-rest schedule. A 1-in-3, 4-h watch schedule causes the person to start watches at the same time every day. This is good for maintaining the synchronization of the body's circadian rhythms with the boat's 24-hour clock.

The physiological aspects of shiftwork scheduling are apparently not the factors that most people use when given a choice of schedule. Sociological factors seem to be perceived as being more important. For example, many nurses prefer 3-crew, 12-h rotating shift systems to 4-crew, 8-h rotating shift systems because the former allow the creation of much longer periods of good quality time off than the latter (Miller, 1992). USCG cutter crew members appeared to prefer slipping from 1 in 3 to 1 in 4 and lower ratios because the longer time between watches for the latter schedules allowed more time for both collateral duties and recovery sleep between watches (Miller *et al.*, 1999). The crew members did not seem to appreciate the physiological advantage of keeping their watch schedule aligned with the day-night cycle: avoiding the malaise associated with circadian rhythm disorder .

Sandquist and colleagues provided a brief review of human fatigue in at-sea work environments (Sandquist *et al.*, 1996). They noted the finding of Rutenfranz and colleagues that watchstanders' average sleep lengths were shorter than those of day workers (Rutenfranz *et al.*, 1988). They also noted the finding by Rutenfranz *et al.* that underway sleep quality was better at night than during the day, and that this finding was consistent with previous sleep studies of shift workers. Sandquist and colleagues also noted the conclusion by Colquhoun (1995) concerning the sleep of maritime watchstanders: "the typical maritime watchstanding schedule leads to incomplete adaptation of physiological circadian rhythms, and that '*the key to such rhythm adaptation lies in the taking of a single, uninterrupted sleep at the same time of day, each day.'"* (Sandquist *et al.*, 1996, pg. 8; emphasis added by Sandquist *et al.*)

Sandquist and colleagues investigated crew fatigue in civilian maritime tankers and freighters using self-reporting by 141 crew members across eight ships (Sandquist *et al.*, 1996). They noted daily sleep times that were too short, very short sleep onset times (indicating excessive sleepiness), and "critically low" alertness levels. They reported, among others, these key findings:

- "Critical levels of fatigue occur between 8 and 21 percent of the time, driven primarily by personnel on the 4-on, 8-off schedule…
- "Mariners sleep an average of 6.6 hours per 24-hour period while on shipboard duty -- this is 1.3 hours less than average sleep duration at home. Sleep debt is known to be cumulative and to reduce performance.
- "Watchstanders generally obtain less total sleep (6.6 hours) than other personnel, and the sleep is of lower quality due to fragmentation and physiologically inappropriate sleep times.
- "Port activities significantly alter the timing of sleep. Frequent changes in sleep timing are known to reduce alertness and performance."
- "The nature and distribution of these findings indicate that the work schedule of the watchstanders is the primary contributor to the fatigue problem." (Sanquist *et al.*, 1996, pg. viii)

Specific characteristics of the traditional maritime watch schedule:

- Cycle length: 24 hours
- Watch length: 4 hours
- Number of watches: 2 in 24 hours

- Longest time off: 8 hours
- Work demand: 8 hours of watch per 24 hours

The 24-hour watch sequence:

Sequence of Watches (4 h)	Scheduled Sleep Time (8 h)	Predicted Best Sleep Times
A: 00-04, 12-16	1600-0000	0430-1030, 1800-2000 (8 h)
B: 04-08, 16-20	N/A	1300-1500, 2100-0300 (8 h)
C: 08-12, 20-00	N/A	1300-1500, 0100-0700 (8 h)

The watch time, scheduled sleep time and predicted best sleep times all sum to 24 h per 72 h and 8 h per 24 h.

The Alternate Watch Schedule

The Alternate schedule was designed to match the work demand of one 4-h watch every 12 h (1 in 3), the traditional maritime watch schedule. It also used the familiar 6-h watch length used presently in submarines. It was a 2-shift, 3-crew solution to the 24/7 work demand, and used an overall time off-to-watch time ratio of 2:1 across 72-h cycles, as in the traditional Maritime schedule. The alternate schedule implements as much as possible the Principles of Chronohygiene suggested by Hildebrandt (1976) and applied in the manual, *Fundamentals of Shiftwork Scheduling,* by Miller (1992). Specifically, the rotating schedule should:

- Minimize the exposure to night work between 24-hour rest periods
- Start each work sequence at noon, when crew members are relatively well rested
- Locate the most error-prone hours, pre-dawn and mid-afternoon, near the start of the sequence, when crew members are relatively well rested
- End each work sequence at noon, allowing adequate recovery sleep to be acquired at night
- Maximize the balance between the number of people available and time spent off from watchstanding; and
- Take advantage of the fact that the human circadian rhythm is longer than 24 hours by rotating forward on the clock.

Specific characteristics of the Alternate schedule:

- Cycle length: 72 hours
- Watch length: 6 hours
- Number of watches: 4 in 72 hours
- Longest time off: 24 hours, noon to noon
- Work demand: 24 hours of watch per 72 hours

The watch sequence:

Sequence of Watches (6 h)	Scheduled Sleep Times	Predicted Best Sleep Times
1200-1800	1900-2300 (4 h)	1900-2300 (4 h)
0000-0600	1300-1700 (4 h)	1300-1700 (4 h)

1800-0000	0100-0500 (4 h)	0100-0500 (4 h)
0600-1200	1400-1600, 2200-0600 (10 h)	1400-1600, 2200-0600 (10 h)
	(24 h off)	(24 h off)

The watch time, scheduled sleep time and predicted best sleep times all sum to 24 h per 72 h, or an average of 8 h per 24 h.

The multi-person (or multi-team) schedule in the figure, below, shows the interactions among people or teams that occurs when the sequences are used. The left-hand part of the schedule shows how people or teams enter the system when the ship gets underway. Note that persons or teams other than A would enter a sequence at a mid-point initially. Also, note that two persons or teams (A and C) start and end their two days of work (and their day off) at 1200, while the other person or team (B) starts and ends their two days of work (and their day off) at 1800.

Two cycles starting at noon for person A, along with the start-ups of the other two watchstanders' schedules. The thicker vertical lines, marked 00, represent midnight. Six days are shown. Each division is six hours. Watches are shown in black.

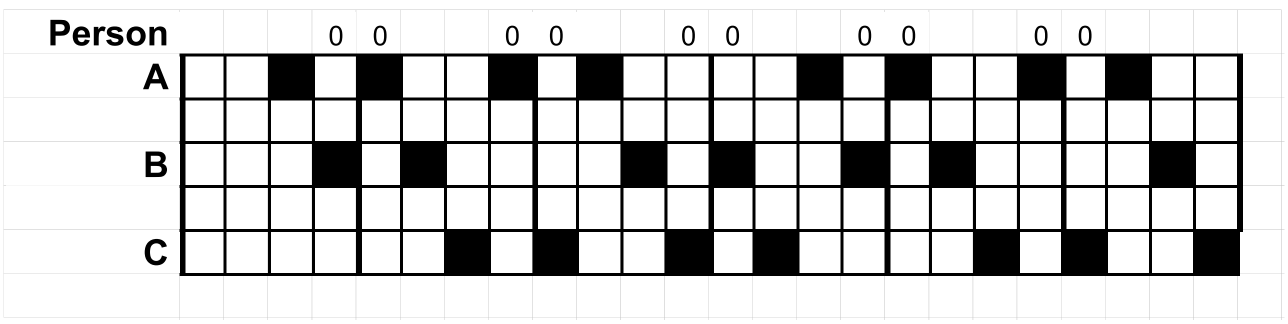

Figure A2. Schematic of the Alternate watch schedule.

Appendix B

Details of Experimental Watchstanding Schedules

Alternate Watch
Experimental Schedule

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
1	12:00	0.0	1					
1	12:30	0.5	1					
1	13:00	1.0	1					
1	13:30	1.5	1					
1	14:00	2.0	1					
1	14:30	2.5	1					
1	15:00	3.0	1					
1	15:30	3.5	1					
1	16:00	4.0	1					
1	16:30	4.5	1					
1	17:00	5.0	1					
1	17:30	5.5	1					x
1	18:00	6.0					1	x
1	18:30	6.5					1	
1	19:00	7.0				1		
1	19:30	7.5				1		
1	20:00	8.0				1		
1	20:30	8.5				1		
1	21:00	9.0				1		
1	21:30	9.5				1		
1	22:00	10.0				1		
1	22:30	10.5				1		
1	23:00	11.0					1	
1	23:30	11.5					1	x
1	0:00	12.0	1					x
1	0:30	12.5	1					
1	1:00	13.0	1					
1	1:30	13.5	1					
1	2:00	14.0	1					
1	2:30	14.5	1					
1	3:00	15.0	1					
1	3:30	15.5	1					
1	4:00	16.0	1					
1	4:30	16.5	1					
1	5:00	17.0	1					
1	5:30	17.5	1					x
1	6:00	18.0					1	x
1	6:30	18.5					1	

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
1	7:00	19.0		1				
1	7:30	19.5		1				
1	8:00	20.0		1				
1	8:30	20.5		1				
1	9:00	21.0		1				
1	9:30	21.5		1				
1	10:00	22.0			1			
1	10:30	22.5			1			
1	11:00	23.0			1			
1	11:30	23.5			1			x
2	12:00	24.0					1	x
2	12:30	24.5					1	
2	13:00	25.0				1		
2	13:30	25.5				1		
2	14:00	26.0				1		
2	14:30	26.5				1		
2	15:00	27.0				1		
2	15:30	27.5				1		
2	16:00	28.0				1		
2	16:30	28.5				1		
2	17:00	29.0					1	
2	17:30	29.5					1	x
2	18:00	30.0	1					x
2	18:30	30.5	1					
2	19:00	31.0	1					
2	19:30	31.5	1					
2	20:00	32.0	1					
2	20:30	32.5	1					
2	21:00	33.0	1					
2	21:30	33.5	1					
2	22:00	34.0	1					
2	22:30	34.5	1					
2	23:00	35.0	1					
2	23:30	35.5	1					x
2	0:00	36.0					1	x
2	0:30	36.5					1	
2	1:00	37.0				1		
2	1:30	37.5				1		
2	2:00	38.0				1		
2	2:30	38.5				1		
2	3:00	39.0				1		
2	3:30	39.5				1		
2	4:00	40.0				1		
2	4:30	40.5				1		
2	5:00	41.0					1	

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
2	5:30	41.5					1	
2	6:00	42.0	1					x
2	6:30	42.5	1					x
2	7:00	43.0	1					
2	7:30	43.5	1					
2	8:00	44.0	1					
2	8:30	44.5	1					
2	9:00	45.0	1					
2	9:30	45.5	1					
2	10:00	46.0	1					
2	10:30	46.5	1					
2	11:00	47.0	1					
2	11:30	47.5	1					x
3	12:00	48.0					1	x
3	12:30	48.5					1	
3	13:00	49.0					1	
3	13:30	49.5					1	
3	14:00	50.0				1		
3	14:30	50.5				1		
3	15:00	51.0				1		
3	15:30	51.5				1		
3	16:00	52.0					1	
3	16:30	52.5					1	
3	17:00	53.0		1				
3	17:30	53.5		1				x
3	18:00	54.0		1				x
3	18:30	54.5		1				
3	19:00	55.0			1			
3	19:30	55.5			1			
3	20:00	56.0			1			
3	20:30	56.5			1			
3	21:00	57.0					1	
3	21:30	57.5					1	
3	22:00	58.0				1		
3	22:30	58.5				1		
3	23:00	59.0				1		
3	23:30	59.5				1		x
3	0:00	60.0				1		x
3	0:30	60.5				1		
3	1:00	61.0				1		
3	1:30	61.5				1		
3	2:00	62.0				1		
3	2:30	62.5				1		
3	3:00	63.0				1		
3	3:30	63.5				1		

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
3	4:00	64.0				1		
3	4:30	64.5				1		
3	5:00	65.0				1		
3	5:30	65.5				1		x
3	6:00	66.0				1		x
3	6:30	66.5				1		
3	7:00	67.0				1		
3	7:30	67.5				1		
3	8:00	68.0					1	
3	8:30	68.5					1	
3	9:00	69.0		1				
3	9:30	69.5		1				
3	10:00	70.0			1			
3	10:30	70.5			1			
3	11:00	71.0			1			
3	11:30	71.5			1			x
		Sum (h)	24	6	6	24	12	

Traditional Maritime Watch
Experimental Schedule

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
1	12:00	0.0	1					
1	12:30	0.5	1					
1	13:00	1.0	1					
1	13:30	1.5	1					
1	14:00	2.0	1					
1	14:30	2.5	1					
1	15:00	3.0	1					
1	15:30	3.5	1					x
1	16:00	4.0				1		x
1	16:30	4.5				1		
1	17:00	5.0				1		
1	17:30	5.5				1		
1	18:00	6.0				1		
1	18:30	6.5				1		
1	19:00	7.0				1		
1	19:30	7.5				1		
1	20:00	8.0				1		
1	20:30	8.5				1		
1	21:00	9.0				1		
1	21:30	9.5				1		
1	22:00	10.0				1		
1	22:30	10.5				1		
1	23:00	11.0				1		
1	23:30	11.5				1		x
1	0:00	12.0	1					x
1	0:30	12.5	1					
1	1:00	13.0	1					
1	1:30	13.5	1					
1	2:00	14.0	1					
1	2:30	14.5	1					
1	3:00	15.0	1					
1	3:30	15.5	1					
1	4:00	16.0					1	
1	4:30	16.5		1				
1	5:00	17.0		1				
1	5:30	17.5		1				
1	6:00	18.0		1				
1	6:30	18.5					1	
1	7:00	19.0					1	
1	7:30	19.5					1	x
1	8:00	20.0					1	x
1	8:30	20.5					1	

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
1	9:00	21.0			1			
1	9:30	21.5			1			
1	10:00	22.0			1			
1	10:30	22.5			1			
1	11:00	23.0					1	
1	11:30	23.5					1	
2	12:00	24.0	1					
2	12:30	24.5	1					
2	13:00	25.0	1					
2	13:30	25.5	1					
2	14:00	26.0	1					
2	14:30	26.5	1					
2	15:00	27.0	1					
2	15:30	27.5	1					x
2	16:00	28.0				1		x
2	16:30	28.5				1		
2	17:00	29.0				1		
2	17:30	29.5				1		
2	18:00	30.0				1		
2	18:30	30.5				1		
2	19:00	31.0				1		
2	19:30	31.5				1		
2	20:00	32.0				1		
2	20:30	32.5				1		
2	21:00	33.0				1		
2	21:30	33.5				1		
2	22:00	34.0				1		
2	22:30	34.5				1		
2	23:00	35.0				1		
2	23:30	35.5				1		x
2	0:00	36.0	1					x
2	0:30	36.5	1					
2	1:00	37.0	1					
2	1:30	37.5	1					
2	2:00	38.0	1					
2	2:30	38.5	1					
2	3:00	39.0	1					
2	3:30	39.5	1					
2	4:00	40.0					1	
2	4:30	40.5		1				
2	5:00	41.0		1				
2	5:30	41.5		1				
2	6:00	42.0		1				
2	6:30	42.5					1	
2	7:00	43.0					1	

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
2	7:30	43.5					1	x
2	8:00	44.0					1	x
2	8:30	44.5					1	
2	9:00	45.0			1			
2	9:30	45.5			1			
2	10:00	46.0			1			
2	10:30	46.5			1			
2	11:00	47.0					1	
2	11:30	47.5					1	
3	12:00	48.0	1					
3	12:30	48.5	1					
3	13:00	49.0	1					
3	13:30	49.5	1					
3	14:00	50.0	1					
3	14:30	50.5	1					
3	15:00	51.0	1					
3	15:30	51.5	1					x
3	16:00	52.0				1		x
3	16:30	52.5				1		
3	17:00	53.0				1		
3	17:30	53.5				1		
3	18:00	54.0				1		
3	18:30	54.5				1		
3	19:00	55.0				1		
3	19:30	55.5				1		
3	20:00	56.0				1		
3	20:30	56.5				1		
3	21:00	57.0				1		
3	21:30	57.5				1		
3	22:00	58.0				1		
3	22:30	58.5				1		
3	23:00	59.0				1		
3	23:30	59.5				1		x
3	0:00	60.0	1					x
3	0:30	60.5	1					
3	1:00	61.0	1					
3	1:30	61.5	1					
3	2:00	62.0	1					
3	2:30	62.5	1					
3	3:00	63.0	1					
3	3:30	63.5	1					
3	4:00	64.0					1	
3	4:30	64.5		1				
3	5:00	65.0		1				
3	5:30	65.5		1				

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
3	6:00	66.0		1				
3	6:30	66.5					1	
3	7:00	67.0					1	
3	7:30	67.5					1	x
3	8:00	68.0					1	x
3	8:30	68.5					1	
3	9:00	69.0			1			
3	9:30	69.5			1			
3	10:00	70.0			1			
3	10:30	70.5			1			
3	11:00	71.0					1	
3	11:30	71.5					1	
		Sum (h)	24	6	6	24	12	

18-Hour Submarine Watch
Experimental Schedule

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
1	12:00	0.0	1					
1	12:30	0.5	1					
1	13:00	1.0	1					
1	13:30	1.5	1					
1	14:00	2.0	1					
1	14:30	2.5	1					
1	15:00	3.0	1					
1	15:30	3.5	1					
1	16:00	4.0	1					
1	16:30	4.5	1					
1	17:00	5.0	1					
1	17:30	5.5	1					x
1	18:00	6.0					1	x
1	18:30	6.5		1				
1	19:00	7.0		1				
1	19:30	7.5		1				
1	20:00	8.0					1	
1	20:30	8.5					1	
1	21:00	9.0				1		
1	21:30	9.5				1		
1	22:00	10.0				1		
1	22:30	10.5				1		
1	23:00	11.0				1		
1	23:30	11.5				1		
1	0:00	12.0				1		N/A
1	0:30	12.5				1		
1	1:00	13.0				1		
1	1:30	13.5				1		
1	2:00	14.0				1		
1	2:30	14.5				1		
1	3:00	15.0					1	
1	3:30	15.5					1	
1	4:00	16.0			1			
1	4:30	16.5			1			
1	5:00	17.0			1			
1	5:30	17.5					1	x
1	6:00	18.0	1					
1	6:30	18.5	1					
1	7:00	19.0	1					
1	7:30	19.5	1					
1	8:00	20.0	1					
1	8:30	20.5	1					

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
1	9:00	21.0	1					
1	9:30	21.5	1					
1	10:00	22.0	1					
1	10:30	22.5	1					
1	11:00	23.0	1					
1	11:30	23.5	1					x
2	12:00	24.0					1	x
2	12:30	24.5		1				
2	13:00	25.0		1				
2	13:30	25.5		1				
2	14:00	26.0					1	
2	14:30	26.5					1	
2	15:00	27.0				1		
2	15:30	27.5				1		
2	16:00	28.0				1		
2	16:30	28.5				1		
2	17:00	29.0				1		
2	17:30	29.5				1		
2	18:00	30.0				1		N/A
2	18:30	30.5				1		
2	19:00	31.0				1		
2	19:30	31.5				1		
2	20:00	32.0				1		
2	20:30	32.5				1		
2	21:00	33.0					1	
2	21:30	33.5					1	
2	22:00	34.0			1			
2	22:30	34.5			1			
2	23:00	35.0			1			
2	23:30	35.5					1	x
2	0:00	36.0	1					
2	0:30	36.5	1					
2	1:00	37.0	1					
2	1:30	37.5	1					
2	2:00	38.0	1					
2	2:30	38.5	1					
2	3:00	39.0	1					
2	3:30	39.5	1					
2	4:00	40.0	1					
2	4:30	40.5	1					
2	5:00	41.0	1					
2	5:30	41.5	1					x
2	6:00	42.0					1	x
2	6:30	42.5		1				
2	7:00	43.0		1				

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
2	7:30	43.5		1				
2	8:00	44.0					1	
2	8:30	44.5					1	
2	9:00	45.0				1		
2	9:30	45.5				1		
2	10:00	46.0				1		
2	10:30	46.5				1		
2	11:00	47.0				1		
2	11:30	47.5				1		
3	12:00	48.0				1		N/A
3	12:30	48.5				1		
3	13:00	49.0				1		
3	13:30	49.5				1		
3	14:00	50.0				1		
3	14:30	50.5				1		
3	15:00	51.0					1	
3	15:30	51.5					1	
3	16:00	52.0			1			
3	16:30	52.5			1			
3	17:00	53.0			1			
3	17:30	53.5					1	x
3	18:00	54.0	1					
3	18:30	54.5	1					
3	19:00	55.0	1					
3	19:30	55.5	1					
3	20:00	56.0	1					
3	20:30	56.5	1					
3	21:00	57.0	1					
3	21:30	57.5	1					
3	22:00	58.0	1					
3	22:30	58.5	1					
3	23:00	59.0	1					
3	23:30	59.5	1					x
3	0:00	60.0					1	x
3	0:30	60.5		1				
3	1:00	61.0		1				
3	1:30	61.5		1				
3	2:00	62.0					1	
3	2:30	62.5					1	
3	3:00	63.0				1		
3	3:30	63.5				1		
3	4:00	64.0				1		
3	4:30	64.5				1		
3	5:00	65.0				1		
3	5:30	65.5				1		

Day	Time of Day	Elapsed Time	Watch	Drill	Train	Sleep	Off	Meal
3	6:00	66.0				1		N/A
3	6:30	66.5				1		
3	7:00	67.0				1		
3	7:30	67.5				1		
3	8:00	68.0				1		
3	8:30	68.5				1		
3	9:00	69.0					1	
3	9:30	69.5					1	
3	10:00	70.0			1			
3	10:30	70.5			1			
3	11:00	71.0			1			
3	11:30	71.5					1	x
		Sum (h)	24	6	6	24	12	

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
10 March 2003	Final	10/01/00 - 09/30/01

4. TITLE AND SUBTITLE

EFFECTS OF THREE WATCHSTANDING SCHEDULES ON SUBMARINER PHYSIOLOGY, PERFORMANCE AND MOOD

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

1) James C. MILLER, Ph.D., CPE,
2) LT Jeff DYCHE, MSC, USNR,
3) Rebecca CARDENAS, MS, and
4) LT Walter CARR, MSC, USNR

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

71845901 (AFRL); 5907 (NSMRL)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

1) AFRL/HEPM
Brooks City-Base, TX 78235

2 & 4) NSMRL
Box 900
Groton, CT 06349-5900

3) NTI, Inc.,
Brooks City-Base, TX 78235

8. PERFORMING ORGANIZATION REPORT NUMBER

NSMRL Technical Report #TR1226

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The general rationale for this study was based upon issues of manpower, especially quality of life and personnel retention, and submarine system performance. Nine male submariners participated as research subjects, two from fast attack submarines and seven from fleet ballistic missile submarines. All participants were submarine qualified and had watchstanding experience on the 18-hour (12-and-6) schedule. The group of participants experienced three watchstanding schedules: the 18-hour Submarine watch schedule; the traditional, Maritime watch schedule; and the Alternative, compressed watch schedule (6-on, 6-off, 6-on, 12-off, 6-on, 6-off-6-on, 24-off). More was learned from the physiological measures employed in this study, especially polysomnography, than from the performance or subjective measures. After combining the results of this study with information available in the research literature about submarine watchstanding schedules dating as far back as 1949, the following schedules were recommended for further study in sea trials: the A Schedule (compressed-6); a fixed, dogged-6 schedule; and the fixed, 8-hour watch schedule.

15. SUBJECT TERMS

Submarine; watchstanding; schedules; sleep; circadian rhythms; 3-shift; 3-team

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U		94	E. Perkins, PublicationsManager
					19b. TELEPHONE NUMBER *(Include area code)* (860) 694-3201

For 24-hour operations in geographically-confined, limited-crew-number situations, consider: the A Schedule (compressed-6); a fixed version of the A Schedule; a fixed, dogged-6 schedule; the fixed, 8-hour watch schedule; or the fixed, 12-and-12 schedule. If work compression and expansion of time off is to be achieved by shortening the watch, then the following schedule should be considered for sea trials and for 24-hour operations in geographically-confined, limited-crew-number situations: the fixed, close watch schedule.